ANNALS OF THE NEW YORK ACADEMY OF SCIENCES
Volume XLVII, Art. 2. Pages 57–240

Editor
Roy Waldo Miner

Consulting Editor
William H. Stein

Associate Editor
Michael Demarest

AMINO ACID ANALYSIS OF PROTEINS

By

William H. Stein, Reginald M. Archibald, Erwin Brand,
R. Keith Cannan, Hans T. Clarke, John T. Edsall, G. L.
Foster, Stanford Moore, David Shemin, Esmond E.
Snell, and Hubert B. Vickery

NEW YORK
Published by the Academy
August 29, 1946

THE NEW YORK ACADEMY OF SCIENCES

(Lyceum of Natural History, 1817–1876)

OFFICERS, 1946

President: Walter H. Bucher

Vice-Presidents: Charles H. Behre, Jr.; Ross F. Nigrelli; Emily T. Burr; George Herzog; Joseph S. Fruton; Athelstan F. Spilhaus

Recording Secretary: Charles M. Breder, Jr.

Corresponding Secretary: H. Herbert Johnson

Treasurer: Harden F. Taylor

Librarian: Erich M. Schlaikjer

Editor: Roy Waldo Miner

Elected Councilors:

1944–1946: Victor K. LaMer; Theodore Shedlovsky; Wilbur G. Malcolm

1945–1947: Ralph H. Cheney; Robert Cushman Murphy; Herbert F. Schwarz

1946–1948: Hans T. Clarke; Charles Glen King; Hans Molitor

Finance Committee: W. Reid Blair (*Chairman*); Marvin D. Thorn; Addison Webb

OFFICERS OF SECTIONS

Geology and Mineralogy

Chairman: Charles H. Behre, Jr. *Secretary:* Arthur N. Strahler

Biology

Chairman: Ross F. Nigrelli. *Secretary:* Charles R. Schroeder

Psychology

Chairman: Emily T. Burr. *Secretary:* S. Stansfeld Sargent

Anthropology

Chairman: George Herzog. *Secretary:* Walter Dyk

Physics and Chemistry

Chairman: Joseph S. Fruton. *Secretary:* Turner Alfrey, Jr.

Oceanography and Meteorology

Chairman: Athelstan F. Spilhaus. *Secretary:* John C. Armstrong

---

The Sections of Geology and Mineralogy, Biology, Psychology, and Anthropology meet in regular session on Monday evenings at 8:00 o'clock from October to May, inclusive. The Sections of Physics and Chemistry, and Oceanography and Meteorology hold two-day conferences at irregular periods. All meetings of the Academy are held at The American Museum of Natural History, Central Park West at 79th Street, New York, N.Y.

ANNALS OF THE NEW YORK ACADEMY OF SCIENCES
VOLUME XLVII, ART. 2. PAGES 57–240
AUGUST 29, 1946

AMINO ACID ANALYSIS OF PROTEINS*

BY

WILLIAM H. STEIN, REGINALD M. ARCHIBALD, ERWIN BRAND, R. KEITH CANNAN, HANS T. CLARKE, JOHN T. EDSALL, G. L. FOSTER, STANFORD MOORE, DAVID SHEMIN, ESMOND E. SNELL, AND HUBERT B. VICKERY

CONTENTS

PAGE
INTRODUCTION. BY WILLIAM H. STEIN 59
THE CONTRIBUTION OF THE ANALYTICAL CHEMIST TO PROTEIN CHEMISTRY. BY HUBERT B. VICKERY 63
THE USE OF SPECIFIC PRECIPITANTS IN THE AMINO ACID ANALYSIS OF PROTEINS. BY WILLIAM H. STEIN AND STANFORD MOORE 95
THE ISOTOPE DILUTION METHOD OF AMINO ACID ANALYSIS. BY DAVID SHEMIN AND G. L. FOSTER 119
CHROMATOGRAPHIC AND ION EXCHANGE METHODS OF AMINO ACID ANALYSIS. BY R. KEITH CANNAN 135
MICROBIOLOGICAL METHODS IN AMINO ACID ANALYSIS. BY ESMOND E. SNELL 161
ENZYMATIC METHODS IN AMINO ACID ANALYSIS. BY REGINALD M. ARCHIBALD 181
AMINO ACID COMPOSITION OF SIMPLE PROTEINS. BY ERWIN BRAND 187
SOME CORRELATIONS BETWEEN PHYSICO-CHEMICAL DATA AND THE AMINO ACID COMPOSITION OF SIMPLE PROTEINS. BY JOHN T. EDSALL 229
CONCLUSION. BY HANS T. CLARKE 237

---

* This series of papers is the result of a Conference on Amino Acid Analysis of Proteins held by the Section of Physics and Chemistry of The New York Academy of Sciences, October 12 and 13, 1945.

Publication made possible through a grant from the Conference Publications Revolving Fund.

gether fitting, therefore, that the methods to be discussed in this conference run the gamut from those which employ the tools of the physicist to those which employ the tools of the bacteriologist and the geneticist.

It is, of course, the immediate goal of all who work on the problem to be able to analyze a protein for its component amino acids with the accuracy and dispatch attending the elementary analysis of simpler compounds. In order to approach this ideal, several conditions must be met. These conditions may be divided into those which apply to methods of analysis, and those which apply to the proteins to be analyzed. Let us discuss, first, the conditions which apply to methods.

The methods required by protein chemists may be differentiated into two types: the primary standard methods and the routine methods. The most important requirement of both types is that they be accurate. The manner of assessing accuracy is different in the two cases, however. Ideally, the accuracy of primary standard methods should rest upon a firm and valid theoretical foundation. If such is the case, it becomes possible readily to establish the limits of error of a method. It would be desirable, also, if primary standard methods could be adapted to the determination of any amino acid in any protein hydrolysate. The primary standard methods, on the other hand, need not, of necessity, be rapid methods, nor need they be micro methods. These latter two useful attributes may be sacrificed in favor of accuracy, if necessary. As an example, the isotope dilution method now appears to possess the characteristics required of a primary standard method.

With routine methods, on the other hand, accuracy is established more on empirical than on theoretical grounds. A routine method should be checked by analysis of known mixtures of amino acids approaching in complexity the composition of a protein hydrolysate. The purity of the amino acids employed in such mixtures should, of course, be established beyond question. In addition, whenever possible, a method should be checked by analysis of a protein which has already been analyzed for the amino acid in question by a primary standard method. Finally, it is important in routine work to have available rapid methods which can be carried out on a micro scale. As examples, certain colorimetric, spectrophotometric, and microbiological methods appear to possess the attributes required of routine methods.

In considering the present status of methods of amino acid analysis, there is apparent need both for more primary standard methods and for quicker, simpler, and more accurate micro methods for the routine determination of many amino acids. Several possibilities for the future development of methods are already apparent. Further work on solubility techniques or on distribution methods (chromatography, counter-current distribution, etc.) might lead to the development of primary standard methods. In the case of routine methods, colorimetry may be extended to include a larger number of amino acids. Many of the objections which have attended the use of colorimetric methods in the past will be removed, once the results of primary standard methods are available, so that the colorimetric procedures may be checked more adequately. Another possibility for development is a logical outgrowth of the microbiological methods. A more intimate knowledge of the intermediary metabolism of both macro- and micro-organisms should permit the preparation of enzymes which may be utilized specifically for the determination of a given amino acid. This eventuality is already in process of realization through the work with arginase, glutaminase, and the specific amino acid decarboxylases.

Just as there is need for primary standard methods of amino acid analysis, there is also need for primary standard proteins to analyze. If the results of amino acid analysis are to be applied to the solution of the problems of protein structure, the choice of standard proteins is crucial, and must be guided by the usages of simple analytical chemistry. Certainly, conclusions as to protein structure are unwarranted, unless the protein investigated conforms to all the known criteria of chemical individuality. Homogeneity in the electrophoresis apparatus, homogeneity in the ultracentrifuge, and adherence to the phase rule requirements for the solubility of a pure substance, are necessary qualifications. Crystallinity is a desirable and esthetically satisfying condition, but it is not, of itself, evidence for purity.

In addition to the minimum requirements just mentioned, additional factors may influence the choice of standard proteins. The advantages of employing a pure protein which is readily available are obvious and require no comment. Proteins possessing an accurately measurable biological activity deserve consideration as standards, since the biological activity of the preparation may be employed both as a descriptive datum of the sample used and as a further criterion of purity. The same advantages attach to the use of proteins which possess precisely determinable non-amino acid moieties, such as iron, copper, zinc, a porphyrin or a carbohydrate. Furthermore, the presence of non-amino acid constituents is of great value in calculating the minimum molecular weight of a protein and in establishing the stoichiometric proportions of the amino acid residues. In view of the many considerations involved, and in the interest of rapid advancement in the field, it would appear profitable to have some concert of opinion as to what proteins might be employed as reference standards.

in this conference, we shall hear of accomplishments that may well serve to give us a feeling of satisfaction and encouragement. This is as it should be, for the more recent developments in this field are of the utmost significance, and nothing but praise should be accorded to those who have penetrated so far. But the fact remains that we are still ignorant of the basic structure of the protein molecule and still unable to account for more than a few of its properties. We are only at the beginning of the phase of explanation and are still busily engaged in fact-gathering.

PROTEIN CHARACTERIZATION

One of the first steps in the elucidation of the structure of an unknown organic compound is its qualitative analysis for the nature of the reactive groups or radicals that endow it with its specific properties. It is not without interest that the first attempt to account for the nature of proteins should have had a somewhat similar motive. Braconnot had observed that, when starch or cellulose from various sources is boiled with acid, the products which result are sweet to the taste, and are undeniably sugars. Accordingly, when, in 1820,[1] he became interested in the composition of certain of the so-called albuminous substances, notably gelatin, he applied a similar decomposition method and was able to prepare a crystalline substance that later became known as glycine, and which, because of its sweet taste, he regarded as a sugar. This was the first amino acid to be obtained in reasonably pure form from a protein after hydrolysis, and Braconnot's observation, although not the first isolation of an amino acid, marks the beginning of our knowledge of the units of which the protein molecule is composed.

The next steps were quantitative ones. Boussingault, in 1836,[2] and Mulder, in the same year,[3] published what were probably the first ultimate analyses of proteins, and Mulder occupied himself with such analyses for many years. From what appeared to him to be regularities in the composition of these substances, he formulated the first hypothesis of their constitution, according to which the substances observed in nature were supposed to be compounds, with differing proportions of sulfur and phosphorus, of a fundamental radical that was given the newly coined name, *protein*. Much has been written concerning the fate of this first attempt to explain the nature of the proteins (*e.g.*, Vickery[4]). All that remains of it today is the name of the radical, with a changed and broadened connotation. But it was from these studies by Mulder, and from the investigations in Liebig's laboratory that they stimulated, that the first attempts to characterize the proteins originated. Mulder supposed that proteins are synthesized only by the plants and that animals secure the protein of their organisms, directly or indirectly, from vegetable sources. Liebig pursued this idea and, as the fruit of his own and of his students' analytical investigations,[5, 6] developed the view that there were only four fundamental substances of this type in nature, namely: albumin, fibrin, casein, and gelatin. These were to be found in both plants and animals.[7]

In justification for the greatly oversimplified nature of this view, it should be remembered that Liebig's only quantitative information was derived from ultimate analysis. He fully recognized the necessity for the study of decomposition products as an additional means for characterization, but was forced to rely entirely upon the results of the purely qualitative hydrolysis and oxidation experiments carried out in his own laboratory. Because of his immense influence upon the chemical thought of the period, Liebig's ideas upon the nature of proteins became a basic premise of protein chemistry for half a century. There was one advantage, however, in spite of the fundamental error. Research was stimulated, inasmuch as comparison of one preparation with another from a different source became necessary in order to establish the limits within which the generalization applied. Thus, Ritthausen, whose investigations of proteins from plant seeds were begun in 1857 (although his first paper[8] did not appear until 1862), soon came to recognize that Liebig's simple scheme had to be broadened, for he found a number of groups of plant proteins that differed both in solubility and in ultimate composition, not only from each other, but from their prototypes of animal origin.

Ritthausen's investigations of proteins extended in unbroken series until his retirement in 1899. His best known publication is his book, *Die Eiweisskörper der Getreidearten, Hülsenfrüchte und Ölsamen,* which appeared in 1872. In this book, Ritthausen, for the first time, squarely faced the problem of protein characterization and differentiation. The title of the book itself reveals the enlargement of his outlook. It had become necessary to group the proteins according to their plant source, inasmuch as the proteins of the cereal grains, the legumes, and the oil-seeds were clearly different in some respects from each other. Nevertheless, Ritthausen adhered as much as possible to a nomenclature founded upon the hypothesis of Liebig. The alcohol-insoluble proteins of wheat and of other cereals were gluten-casein, the alcohol-soluble proteins were a mixture of plant gelatin (*Pflanzenleim*), gluten-fibrin, and what he termed mucedin. These were three fractions separated by various methods from what is today termed gliadin, mucedin being the most soluble in dilute alcohol. From the seeds of buckwheat, he obtained preparations that he called plant casein, Possibly, we may have, in the course of the discussion here, some expression of views on this matter.

Another point relative to the amino acid analysis of proteins is worthy of consideration. Strictly speaking, the title of this conference is a misnomer, for there doubtless will be little discussion concerning the amino acid analysis of proteins. Rather will interest be centered on the amino acid analysis of protein hydrolysates. Before it is legitimate to apply the results obtained on protein hydrolysates to the problems of protein structure, it is first necessary to establish that the amino acid composition of a given hydrolysate is a true reflection of the amino acid composition of the protein from which it is derived. Or, to state the problem more directly, it is necessary to prove that decomposition of amino acids, either at the peptide stage or at the free amino acid stage, does not occur during hydrolysis. Furthermore, when the use of methods possessing antipodal specificity is contemplated, it is also necessary to preclude the occurrence of racemization during hydrolysis.

The necessity for a critical examination of this question is attested by evidence in the literature which indicates that decomposition or racemization of amino acids may occur during hydrolysis. It is well known, for example, that amino acids combined in peptide linkage are more susceptible to decomposition or racemization by hot acids and alkalis than they are when free. Indeed, the behavior of peptides during hydrolysis poses a dilemma, the imminent solution of which does not appear likely.

The question of the decomposition or racemization of free amino acids during hydrolysis is, however, susceptible to immediate attack. Work on this problem has been hampered by a dearth of methods. In order to attack the question, simple, rapid, and accurate methods for the determination of numerous amino acids are essential. With the advent of the newer methods, particularly the microbiological and colorimetric procedures, an approach to this problem appears feasible. In fact, the time is at hand when a re-examination of the whole question of methods of hydrolysis may be in order. Milder methods are needed, as is a systematic investigation of the conditions of hydrolysis which are optimum for the determination of each amino acid.

In surveying the field to be covered by this conference, then, notable recent achievements are apparent. Much, however, still remains to be done. The application of the methods of today and tomorrow to the solution of the manifold problems of protein chemistry and physiology doubtless will become a fascinating and fruitful field for research. This conference will have served its purpose well, if it hastens by even a little the time when we shall know more about proteins.

THE CONTRIBUTION OF THE ANALYTICAL CHEMIST TO PROTEIN CHEMISTRY

By Hubert Bradford Vickery

From the Connecticut Agricultural Experiment Station, New Haven, Connecticut

The present conference on methods for the determination of the amino acids that are produced by the hydrolysis of proteins has, as its primary object, the discussion of analytical techniques that have been developed in recent years. As an introduction, it seems desirable to consider the reasons why the analytical chemistry of the amino acids has attracted so much attention in the past, and to attempt to assess the value to science of the information that has been provided.

The point of view of those who have concerned themselves with the fundamental problem of the composition of the protein molecule has undergone marked changes with the passage of the years. The earliest analyses of proteins were carried out in order to characterize and differentiate protein preparations from various sources. Subsequently, there developed an interest in the composition of the protein molecule as a whole: in particular, the desire to see if it is composed exclusively of units that yield amino acids on hydrolysis, or if units of some other kind are involved to any significant extent. In recent years, although both of the original motives have maintained their place in the thoughts of protein chemists, there has arisen a conviction that not only the chemical properties of the proteins, but also the physical properties, or at least many of them, can be assigned a rational explanation in terms of the amino acid composition, if this is sufficiently well known and can be adequately interpreted.

Finally, still far on the horizon, but becoming ever brighter, is the hope that we shall ultimately be able to formulate a comprehensive and convincing theory of the structure of the protein molecule, a single brief statement that will shed light alike upon the chemical, physical, enzymological, immunological, and, in appropriate cases, the hormonal properties of these substances, and indicate the way for all subsequent research upon them.

It is my purpose to trace briefly the reasons for these changes in the ambition of the protein chemist, but the point that I wish to emphasize at the start is that, in spite of the devotion and zeal with which the many problems of protein analysis have been attacked for more than a century, we have not yet attained a final solution. In the papers that are to follow and from several species of beans, he prepared a kind of plant casein which received the name of legumin. Lupine seeds and almonds were found to yield preparations that differed from legumin, although they had the fundamental properties of a plant casein. This product was named conglutin.

The differentiation and characterization of these protein materials was by no means easy. Ritthausen, in fact, failed to make entirely clear the grounds upon which he relied. He depended upon solubility primarily, but he used the terms albumin (*Eiweiss*), plant casein, or plant gelatin as group designations. Under the heading plant casein, he placed legumin and conglutin, as well as the less clearly differentiated gluten casein, which is now called glutenin. In justice to Ritthausen, it is only fair to state, with respect to this last product, that we have today merely given another name to a protein system that is still scarcely at all understood. Plant gelatin was a group designation employed for the alcohol-soluble proteins of cereal seeds, gliadin, mucedin, and gluten-fibrin, which were thought to be present in different proportions in the various cereal grains. We define gliadin today as the system of alcohol-soluble, protein components of wheat seeds, and have dropped Ritthausen's other terms in favor of specific designations (such as zein for the alcohol-soluble protein of maize) or, occasionally, of Osborne's general term, prolamins, for proteins of this type.

The other fundamental criterion was the ultimate composition. This, however, could not be applied rigorously. The three main kinds of plant caseins, for example, showed differences in carbon, hydrogen, and sulfur, but within each group there were variations. Ritthausen (page 231[9]) stated that it was necessary to assume that these variations arose from the differences in the relative proportions of the component parts of the proteins. What he meant by this is clear from his discussion of their decomposition products. He pointed out that the three kinds of plant casein, legumin, conglutin, and gluten-casein, differ in the relative proportions of aspartic and glutamic acids that can be obtained from them after complete hydrolysis, and, accordingly, he concluded that the three substances that had been designated plant casein are not identical with each other. Thus, in 1872, Ritthausen was already making use of the amino acid composition as a means to differentiate proteins of somewhat similar properties, although, to be sure, in only a tentative manner.

Because of the historical importance of this approach to the problem, it may be well to digress, in order to point out the sequence of events. Ritthausen discovered glutamic acid in 1866,[10] and identified aspartic acid as one of the decomposition products of proteins in 1869.[11] In 1871, he and Kreusler[12] published a brief paper in which their attempts to determine the quantities of these two acids yielded by different proteins were summarized. The table in this paper (TABLE 1) is, so far as I have been able to discover, the first table of the amino acid composition of proteins to appear in print. They point out that, although neither aspartic nor glutamic acids can be determined with much accuracy, the figures serve to illustrate the relations among the proteins mentioned. Furthermore, differences so great as those observed in the yield of glutamic acid can scarcely be regarded as the result of chance conditions affecting the success of the isolation, but must mean that the proteins themselves differ from each other. Subsequent research has shown that they were entirely correct in this conclusion.

TABLE 1
(Ritthausen and Kreusler, 1871)

Proteinkörper 100 Gew.-Th. gaben:	Asparagin-	Glutaminsäure
1) Mucedin	nicht bestimmt	25 p.c.
2) Maisfibrin	1.4 p.c.	10.0 "
3) Gemenge von Pflanzenleim, Mucedin und Fibrin	1.1 "	8.8 "
4) Glutencasein	0.33 "	5.3 "
5) Conglutin (Lupinen)	2.0 "	3–5 "
6) Legumin (aus Saubohnen)	3.5 "	1.5 "

The table was reprinted in *Die Eiweisskörper* (page 222[9]), with slight modifications in the designations of the preparations analyzed and with further discussion. The book also contains the analytical data upon which the identification of the products of hydrolysis rests.

Ritthausen's emphasis upon the importance of the decomposition products was, of course, strictly in the tradition that originated with Braconnot and was carried on by Liebig. Braconnot had prepared glycine from gelatin, and a white crystalline mass, that he designated leucine, from other proteins. Liebig had encountered tyrosine as well as leucine among the products of the fusion of casein and of other proteins with potassium hydroxide,[13] and two of his students[14, 15] had obtained both of these products after acid hydrolysis. Another student[16] had studied the products of oxidation and had identified a series of lower fatty acids and aldehydes, as well as benzaldehyde. None of this work, however, had been carried out quantitatively. It was Ritthausen who first attempted to establish quantitative relationships.

The significance of this approach was at once grasped by Schulze, who was concerned with the metabolism of the proteins of seeds during the process of sprouting. He applied it to show that the composition of the extract obtained from the seedlings was entirely different, with respect to its amino acid composition, from that to be expected from the composition of the seed protein, if no metabolic process other than hydrolysis had occurred.[17] Schulze, together with Barbieri, attempted the amino acid analysis of the main proteins of the seeds that he was studying, and was perspicacious enough to employ the new method of hydrolysis with hydrochloric acid and stannous chloride, introduced by Hlasiwetz and Habermann,[18] in 1873. The results (TABLE 2) were given with reservations as to accuracy, because of the well-recognized difficulties of fractional crystallization, but the values for conglutin were held to confirm, in general, the results of Ritthausen. Thus, at the very dawn of amino acid analytical chemistry, a practical application of the results was made to assist in the solution of a problem in metabolism, and the custom of checking and verifying both the methods and the data was initiated.

TABLE 2
(Schulze, 1880)

	Cucurbit seed protein	Lupine seed protein (Conglutin)
	per cent	*per cent*
Leucine	20	10
Tyrosine	2	2
Aspartic acid	2.5	1.5
Glutamic acid	3.4	6

Ritthausen's extensive studies of the seed proteins, careful as they were, left many problems of characterization unsolved. Aside from ultimate analysis, he knew of no analytical methods accurate enough to provide trustworthy information, and the narrow range in composition of the proteins, with respect to carbon, hydrogen, nitrogen, and even sulfur, made clear differentiation impossible. Osborne's investigations, which were begun in 1889 and extended for many years, shed much light upon the behavior, especially the solubility, of this group of proteins. Yet, throughout the last decade of the century, there was no alternative method of characterization, aside from precipitation reactions and coagulation temperature. All that could be done was to prepare these substances with the utmost care in specification of conditions (with respect to salt concentration, temperature, acidity, and nature of the solvent) under which they were obtained,* and to record their ultimate composition. As an illustration of the stage which Osborne's investigations had reached in 1896 and of the point of view he had adopted, as well as of the application of the available methods of characterization, TABLE 3 is presented. This is taken, with some simplification, from a paper by Osborne and Campbell[19] entitled "Conglutin and Vitellin." The name, conglutin, was used by Ritthausen to designate certain of the so-called plant caseins, while vitellin was a term employed by Weyl[20] for seed proteins which, like the protein of egg yolk, were soluble in 10 per cent salt solution, but were not precipitated by saturation of the solution with salt, and which coagulated when heated to 75°. Osborne's purpose was to reexamine all of the proteins that had been de-

---

* Osborne's work on the preparation of proteins was completed long before the development of the physical-chemical theory of acidity and the concept of solubility as a function of ionic strength. Nevertheless, his experimental handling of these factors, although empirical, was masterly.

TABLE 3

CONGLUTIN AND VITELLIN
(Osborne and Campbell, 1896)

	Edestin	Amandin	Corylin	Excelsin	Avenalin	Conglutin
	per cent	*per cent*	*per cent*	*per cent*	*per cent*	*per cent*
Carbon	51.65	51.30	50.72	52.18	52.18	51.00
Hydrogen	6.89	6.90	6.86	6.92	7.05	6.90
Nitrogen	18.75	19.32	19.17	18.30	17.90	17.99
Sulfur	0.85	0.44	0.83	1.06	0.53	0.40
Oxygen	21.86	22.04	22.42	21.54	22.34	23.71
Precipitation reactions						
Salt solution saturated with:						
Sodium chloride	None	None	None	None	Complete	None
Magnesium sulfate	Complete	Partial	Partial	Slight	Complete	None
Mercuric chloride	Ppt.	None	Ppt.	None	Ppt.	None
10% solution of protein in 10% NaCl diluted with 1 volume H_2O	Ppt.	None	Ppt.	Slight	Ppt.	None
Heat coagulation						
Turbidity	88°	75°	80°	70°	None	Trace at 99°
Flocculation	95°	80°	99°	84°	None	
Distribution						
In seeds of:	Hemp Castor bean Squash Flax Cotton Wheat Rye Barley Maize Cocoanut	Almond Peach	Walnut Filbert	Brazil-nut	Oat	Lupine

scribed under these two names and to see if they were indeed alike. By paying careful attention to the precipitation reactions and coagulation temperatures, as well as to the ultimate analysis, he was able to set up six different groups of proteins under new designations, three of these having been found only in a single seed. Many of the preparations were obtained in crystalline form, and all had been purified by repeated separation from solution. This was done by diminishing the ionic strength through dialysis, or by dilution at an elevated temperature and allowing the warm solution to cool slowly.

The outstanding problem presented by this table is the nature of the preparations designated as edestin. This was a term coined by Osborne for seed globulins of the composition and behavior shown, and, at the time, he seems to have entertained little doubt that the preparations from different sources actually represented the same substance. Nevertheless, he was well aware that the evidence was inconclusive, and, some years later, he returned to the problem of protein differentiation. He was now in possession of methods adequate for the task. Lysine had been discovered by Drechsel,[21] in 1889; arginine had been shown to be a component of proteins by Hedin,[22] in 1895; and histidine had been discovered, simultaneously, by Hedin[23] and by Kossel,[24] in 1896. Kossel and Kutscher[25] had devised an accurate method for the quantitative determination of the three basic amino acids, while Hausmann,[26] a student in Hofmeister's laboratory, had developed a simple and rapid method for the determination of the proportion of the nitrogen of the protein that was present in the basic amino acids. Osborne recognized that the Kossel method was alone capable of providing final proof of differences in the composition of these proteins. However, since large quantities of each preparation would be required, as well as a great deal of time to complete the analyses, he, together with Harris,[27] in 1903, made a preliminary survey of the composition of the numerous preparations of seed proteins that had been accumulated over the years. From the comprehensive paper that resulted, only one table will be reproduced, in abbreviated form, since it led to a partial solution of the edestin problem. This is shown in TABLE 4. From these data, Osborne concluded that the globulins of hempseed, cottonseed, and castor bean were probably indistinguishable from each other. However, the cottonseed globulin gives a strong Molisch reaction, while those of the hempseed and castor bean give none. Thus, only the hempseed and the castor bean globulins still remained indistinguishable, and, since the many other analyses reported showed that there were, in general, close resemblances in composition only in the cases of the similar proteins derived from closely related plant species, he felt it desirable to restrict the name, edestin, to the globulin from hempseed. The other proteins in the table were clearly different from each other in composition.

Convincing evidence was presented, in 1908, in a paper[28] which reported the results of the application of the Kossel method for the determination of the bases. No two of the twenty-six proteins examined resembled each other sufficiently closely to be confused. Furthermore, when the proteins were arranged in the order of their yield of arginine, they fell into three fairly readily distinguishable groups, the proteins of the oil seeds being at the top, those of the legumes in an intermediate position, and those of the cereals at the bottom. Osborne pointed out that Ritthausen had recognized a similar grouping of proteins as a result of his studies of their general properties, and had considered it of sufficient significance to employ in the title of his book.

TABLE 4
EDESTIN
RESULTS OF NITROGEN DISTRIBUTION ANALYSIS BY THE HAUSMANN METHOD
(Osborne and Harris, 1903)

Source	Ammonia N	Basic N	Non-basic N	Humin N	Total N
	per cent	*per cent*	*per cent*	*per cent*	*per cent*
Hempseed	1.86	5.98	10.68	0.13	18.64
Cottonseed	1.92	5.71	11.01		18.64
Castor bean	1.96	5.64	11.03	0.12	18.75
Cocoanut	1.38	6.11	10.79	0.20	18.48
Flaxseed	1.89	4.68	11.64	0.27	
Squash seed	1.36	5.93	11.04	0.18	18.51
Sunflower seed	2.58	4.21	11.55	0.24	18.58
Wheat kernel	1.49	6.66	10.04	0.20	18.39

It was apparently Ritthausen who first grasped the complexity of the problem presented by the differentiation of proteins, and it was he who first proposed the analytical determination of amino acids as a means to this end. His own attempts were necessarily unconvincing, since they were restricted to the determination of only two amino acids — substances that were his own discoveries and for which adequate analytical methods have been developed only in the past few years. Nevertheless, he pointed out the way. Osborne's work clarified much that was left vague and uncertain by the great German investigator, and it is to him that we owe the application of rigid methods of comparison in a systematic manner. He made use of every technique available to him; each of the recognized color reactions was studied and applied; such physical properties as specific rotation, heat of combustion, the solubility in salt solutions, and the precipitation limits with ammonium sulfate were examined; and amino acid analyses were carried out in the effort to discern differences, since he had become convinced that the demonstration of a difference was far more significant than the observation of a similarity. Ultimately, the problem was completely cleared up by the application, with Wells,[29] of immunological reactions which showed that there were only two or three cases of proteins from closely related botanical species that could not be demonstrated to be different from each other in one respect or another.

In following through the development of the methods of characterization as applied to proteins of plant origin, reference has been omitted to one other striking example of a clean-cut differentiation between allied proteins by strictly analytical methods. The fact that the hemoglobin of blood could be obtained in crystalline form early attracted attention,[30, 31] and methods for the preparation of highly purified specimens were developed by Hoppe-Seyler.[32] In 1886, Zinoffsky,[33] and, two years later, Jaquet,[34] undertook a series of analytical studies of repeatedly crystallized specimens of the hemoglobin of the horse, the dog, and the fowl, especial attention being given to the content of iron and sulfur. A summary of their results is shown in TABLE 5. This work was carried out with the utmost care, and many of the determinations were separately confirmed by different analytical methods by Bunge, under whose direction the investigation was conducted. Although Zinoffsky's nitrogen value in horse hemoglobin is undoubtedly in error, the accuracy these workers attained in the determination of iron and sulfur has rarely been equalled and probably has never been excelled. Their values for iron, in particular, are still quoted as authoritative. Eyebrows may be raised at the use of four figures in the percentage of iron and sulfur, but careful study of the details of the data

TABLE 5

COMPOSITION OF HEMOGLOBIN
(Zinoffsky, 1886; Jaquet, 1890)

	Horse	Dog	Fowl
	per cent	*per cent*	*per cent*
Carbon	51.15	54.57	52.47
Hydrogen	6.76	7.22	7.19
Nitrogen	17.94	16.38	16.45
Sulfur	0.3899	0.568	0.8586
Iron	0.3351	0.336	0.3353
Oxygen	23.42	20.93	22.5
Phosphorus			0.1973

in the original publications will show that the third figure is fully significant. These analyses permitted the definite conclusion that the respective hemoglobins are different substances. This point was further emphasized by Gamgee[35] in the light of these and of other, later analyses which showed that the ratio of iron to sulfur varies within wide limits, according to species.

THE COMPOSITION OF THE PROTEIN MOLECULE

In 1901, on June 1, Kossel delivered a lecture before the German Chemical Society on the contemporary position of protein chemistry.[36] This lecture is of great historical interest, for it thoroughly summarized the accomplishments in the field up to that date, though presented from the point of view current just before the revolution in thought brought about by the peptide hypothesis of Fischer and Hofmeister, and by the introduction of the ester distillation method for the separation of the mono-amino acids. As a matter of fact, Fischer's manuscript, "On the hydrolysis of casein by hydrochloric acid,"[37] the paper in which the ester method was first described, together with the discovery of proline and valine as products of protein hydrolysis, was submitted to Kossel, as editor of Hoppe-Seyler's *Zeitschrift,* on July 5th of that year. Kossel, therefore, was able to introduce a statement about proline into the proof of the printed lecture. This statement he carefully enclosed in brackets, and a footnote was added to indicate its later origin.

In the lecture, Kossel reviewed the long list of substances that had been reported as products of the decomposition of proteins, and summarized the most significant of these as follows (TABLE 6): (1) The urea-forming group of arginine; (2) the group of diamino acids, diamino-acetic, diamino-valeric and diamino-caproic acids; (3) the group of mono-amino acids, which included glycine, an amino-valeric acid, an amino-caproic acid,

TABLE 6

SUMMARY OF PROTEIN COMPOSITION
(Kossel, 1901)

1. The urea-forming group of arginine.
2. The diamino acids, diamino-acetic, diamino-valeric, and diamino-caproic acids.
3. The mono-amino acids, glycine, amino-valeric, amino-caproic, amino-thiolactic and its corresponding disulfide, an unknown sulfur-containing substance, two amino-propionic acids substituted with the benzene ring, a skatole- and indole-forming substance, and aspartic and glutamic acids.
4. Pyrrolidine-carboxylic acid (added in proof) and the furfural, the ammonia, and the humin-forming substances.

leucine, and an amino-thiolactic acid, cysteine, and its disulfide, together with a still unknown sulfur-containing substance, two amino-propionic acids substituted with the benzene ring (phenylalanine and tyrosine), and the widely distributed aspartic and glutamic acids. In addition, he mentioned a skatole- and indole-forming substance that probably belonged to this group. (4) Finally, there was a group of products of protein decomposition, the relation of which to these better-defined substances was not clear; this included pyrrolidine-carboxylic acid, furfural, ammonia and the humin-yielding substance. Of these, Kossel felt that furfural doubtless originates from a special group in the protein molecule.

This list contains remarkably few errors. Subsequent research has discarded diamino-acetic acid entirely and has changed the emphasis upon diamino-valeric acid or ornithine, but it must be remembered that the exact structure of arginine was still a matter of speculation at this time, and arginine was chiefly interesting because of its relationship to urea. It is perhaps remarkable that Kossel disregarded the early suggestions of Ritthausen and of Hlasiwetz concerning the relationship of the ammonia to the dicarboxylic amino acids. Their views were, however, consistently overlooked, until revived by Osborne, Leavenworth, and Brautlecht, in 1908.[28] Even more remarkable, perhaps, is the modest omission of Kossel's own discovery of histidine from the list. It is referred to briefly in the text as a basic substance of unknown constitution which was thought to be widely distributed as a component of proteins. The omission of alanine and serine may perhaps occasion comment, but it was the ester method of Fischer that demonstrated that these substances are common components of protein hydrolysates. The isolation of tryptophane by Hopkins and Cole,[38] and the demonstration that it is the skatole-forming substance, were not described until several months after Kossel's lecture was published, although the name, "tryptophane," had been in use since 1890, as a designation for the group responsible for certain color reactions of proteins.

Kossel pointed out that the decomposition products mentioned represented experimentally a large number of different proteins, but that, nevertheless, most proteins yielded many of them. Accordingly, the protein molecule, doubtless, must be extraordinarily complicated, and the development of a structure formula promised to be an assignment of the greatest difficulty, even if preparations of demonstrated purity could be secured — in itself an extremely difficult task. Kossel also emphasized the wide differences that had been found in the analytical composition of proteins with respect to amino acids, especially in the cases of arginine and of glutamic acid. His own work upon the protamines had revealed the existence of protein-like substances, some of which yielded more than 80 per cent of arginine, while such plant proteins as zein yielded as little as 1.8 per cent. Thus, the proteins comprised a group of substances of the widest possible variety in composition. In the final paragraph, he summed up the situation as follows: "It has been customary to think of protein as a substance of definite fixed properties, a sort of 'ideal protein,' much as Goethe thought of an '*Urpflanze*' or 'ideal plant.' Those protein substances that did not correspond to this ideal have been supposed to be defective and have been placed in a lower group, such as the albuminoids. This point of view can no longer be maintained. It is a necessity in present day science to regard each organic substance as a member of an evolving series, a necessity that has its most striking expression in the direction taken by phylogenetic and ontogenetic research. We must not, therefore, consider one complicated protein molecule to be representative of all,* but must seek to find a system of proteins which, progressing from the simplest to the most complex, reveals to us the innermost character of these many-sided substances."

It cannot be claimed that we have developed, even today, a system of proteins in the sense that Kossel implied, although much progress has been made. Nevertheless, the climactic change that occurred within a few years after his lecture was delivered was a result of the longest single stride that has even yet been taken in protein chemistry. This was the concept of union between the amino acids in the protein molecule by what are, in effect, substituted amide bonds; namely, the peptide bond hypothesis of structure. The idea is implicit in Fischer's earliest papers on the peptides,[40] and the possibility that this type of combination could be brought about synthetically had been frequently pointed out by earlier workers, who had prepared complex anhydrides of several amino acids (*e.g.*, Schaal,[41] Schiff[42]), and especially by Curtius who had prepared glycine anhydride in 1888.[43] A clear statement of the hypothesis as it applies to the protein molecule was, however, first made by Hofmeister,[44] in 1902, and greatly stimulated the attempts which soon followed to obtain complete analyses of the protein molecule.

There had, indeed, been earlier efforts to account for the whole of the protein in terms of its decomposition products. One of the most ambitious of these was made by Hlasiwetz and Habermann, in 1873,[18] in a paper that is one of the great classics of protein chemistry. As a result of their fractionation of the products of hydrolysis of casein, they stated that this protein is composed entirely of glutamic acid, aspartic acid, leucine, tyrosine, and ammonia. However, the only item of quantitative data that they gave in support of this assertion was the yield of glutamic acid, of which they claimed to have isolated 29 per cent. It is obvious from this single figure that they were content to weigh extremely crude material.* Being themselves so uncertain regarding the yields of the other decomposition products, they thought it better to omit all figures, with the promise to return to the quantitative aspect of the problem at a later date. Hlasiwetz' death, in 1875, prevented this.

---

* This statement is of importance today. A recent monograph on protein analysis[39] advocates the computation of all protein analytical results upon the basis of a hypothetical substance that contains 16 per cent of nitrogen. Such a substance is exactly what Kossel meant by the term "ideal protein," and the proposal thus represents a point of view to which he saw serious objections as long as 45 years ago.

Fischer's introduction of the ester method for the separation of the mono-amino acids, together with Kossel and Kutscher's method for the determination of the basic amino acids and the direct isolation procedures for glutamic acid, cystine, and tyrosine, made it feasible, during the early years of the present century, to attempt the complete analysis of proteins. Of the innumerable papers that appeared during the next few years, only a few have lasting value. Unfortunately, the field became a happy hunting-ground for the graduate student in search of a thesis problem, and the results, as a whole, were thoroughly discouraging. Nevertheless, Fischer himself[45] carried out a few analyses, notably one of the silk of a large Madagascar spider (TABLE 7), the results of which, expressed as amino acids, added up to over 90 per cent and accounted, according to his method of calculation, for about 74 per cent of the molecule, when allowance for the water of hydrolysis was made.

An even more striking result was obtained by Kossel and Dakin,[46] in 1904. Salmine from the Rhine salmon was found by the ester distillation method to yield only valine, proline, and serine, in addition to arginine. By taking advantage of this qualitative simplicity, they devised a fractionation method that led to the extraordinarily complete results shown in TABLE 8. In 1908, Taylor,[47] in a rarely quoted paper, applied the ester method, together with the Kossel method, to the salmine from a Pacific coast salmon and obtained figures which are also shown in TABLE 8. Kossel later[48] demonstrated that the proportion of the protamine nitrogen present as arginine nitrogen was nearly identical in these two kinds of salmine, but whether or not they are in fact identical substances does not even yet appear to be known.

Osborne, in this country, began his studies of the analysis of proteins by the ester method in 1906,[49] the first paper dealing with the proteins of wheat. The data of this set of analyses are shown in TABLE 9, and merit attention as an example of what the original ester method was capable of achieving. For the gliadin and glutenin analyses, almost a kilogram of protein was taken, and the glutamic acid was separated, as completely as possible, as the hydrochloride. The mother liquors were then combined, and the amino acids they contained were esterified. After liberation of the esters and extraction with ether, the mass of salts was extracted with alcoholic hydrochloric acid to recover amino acid material, and the process of esterification was repeated. The second lot of esters was kept separate,

---

* Present-day methods show that casein yields about 22 per cent of glutamic acid. On the other hand, it is not impossible that Hlasiwetz and Habermann calculated the yield as glutamic acid hydrochloride, in which case they were not very much in error. Nevertheless, later research has shown that it is difficult to isolate more than half of the glutamic acid of casein by the method they used.

TABLE 7

AMINO ACIDS OF SPIDER SILK (*Naphila madagascariensis*)
(Fischer, 1907)

"The total value of 91.5 per cent is diminished to 74 per cent if the water added by hydrolysis is deducted. The amount assumed for the first five amino acids was 1 mole and for the glutamic acid 2 moles."

100 parts of dry silk yielded:	
Glycine	35.13
Alanine	23.4
Leucine	1.76
Proline	3.68
Tyrosine	8.2
Glutamic acid	11.7
Diamino acids as arginine*	5.24
Ammonia	1.16
Fatty acids	0.66
	90.93
Ash	0.59

* The sum of the basic amino acids was calculated from the nitrogen precipitated by phosphotungstic acid.

TABLE 8

AMINO ACIDS OF SALMINE
(Kossel and Dakin, 1904; Taylor, 1908)

	Rhine salmon *Salmo salar*	Pacific coast salmon *Oncorhynchus tschawytscha*
	per cent	*per cent*
Arginine	87.4	91.73
Valine	4.3	5.35
Proline	11.0	10.83
Serine	7.8	8.70
	110.5	116.61

Table 9

Analysis of Wheat Proteins
(Osborne and Clapp, 1906)

	Gliadin	Glutenin	Leucosin
	per cent	*per cent*	*per cent*
Glycocoll	...	0.89	0.94
Alanine	2.00	4.65	4.45
Amino-valerianic acid	0.21	0.24	0.18
Leucine	5.61	5.95	11.34
α-Proline	7.06	4.23	3.18
Phenylalanine	2.35	1.97	3.83
Aspartic acid	0.58	0.91	3.35
Glutaminic acid	37.33	23.42	6.73
Serine	0.13	0.74	...
Tyrosine	1.20	4.25	3.34
Cystine	0.45	0.02	...
Oxyproline	...	...	...
Lysine	0.00	1.92	2.75
Histidine	0.61	1.76	2.83
Arginine	3.16	4.72	5.94
Ammonia	5.11	4.01	1.41
Tryptophane	present	present	present
Total	65.80	59.68	50.32

and a third esterification of material recovered from the salts was carried out, the esters being combined with those from the second process. The distillation of these two lots of esters yielded fractions that were saponified, and the amino acids were separated, as completely as possible, by fractional crystallization. The preparations were identified by ultimate analysis for carbon and hydrogen, and often were further characterized by specific rotation and the preparation of derivatives. Occasionally, small intermediate fractions were assumed to be mixtures of two adjacent amino acids, and the composition of the mixture was calculated from the carbon content. The data for the first nine amino acids in the table were obtained from these results. Cystine was determined in a separate hydrolysate of 300 gm. of protein, a short hydrolysis time being employed in order to minimize racemization. The glutamic acid was crystallized out and the mother liquors were neutralized. From the mixture of amino acids that then separated, the cystine was precipitated with mercuric sulfate and isolated by crystallization. Tyrosine was isolated from still another hydrolysate by direct crystallization. The bases were determined by the Kossel and Patten method.[50]

Analyses of this kind were conducted in Osborne's laboratory almost continuously for the next six years. He repeatedly analyzed the most important food proteins such as gliadin, zein, and casein, always with improvement in the results. It soon became evident, however, that the technical difficulties were so great that the ideal of a complete analysis, for which high hopes had been entertained at the start, was unattainable. In 1910, Osborne and Jones[51] considered the sources of loss in protein analysis, and carried out an analysis of a mixture of 328 gm. of pure amino acids, in order to obtain information upon the accuracy of the method. The results are shown in TABLE 10. The figures in the last column revealed at once how little confidence could be placed in the data, and turned attention toward the necessity, not only for improvements in the original technique, but also for entirely new methods of analysis.

Table 10

Analysis of a Known Mixture of Amino Acids
(Osborne and Jones, 1910)

		Recovered			Recovered		
Amino acids taken		Esters I	Esters II	Total	Esters I	Esters II	Total
	gm.	*gm.*	*gm.*	*gm.*	*per cent*	*per cent*	*per cent*
Alanine	16.00	7.34	0.00	7.34	45.88	0.00	45.88
Valine	4.00	1.40	0.24	1.64	35.00	6.00	41.00
Leucine	85.00	53.18	16.18	69.36	62.56	19.04	81.60
Proline	31.00	19.54	3.02	22.56	63.03	9.74	72.77
Phenylalanine	26.00	15.64	2.43	18.07	60.16	9.35	69.51
Aspartic acid	6.00	2.55	0.00	2.55	42.50	0.00	42.50
Glutaminic acid	120.20	50.11	33.15	83.26	41.76	27.63	69.39
Tyrosine	16.00	7.56	0.41	7.97	47.25	2.56	49.81
Arginine	5.50	3.57	0.00	3.57	64.91		64.91
Histidine	2.04	0.80	0.00	0.80	39.21		39.21
Serine	2.00	0.00	0.00	0.00	0.00		0.00
Ammonia	14.31						
Total	328.05	161.69	55.43	217.12			66.17

Notable among the improvements that were suggested in the technique of the ester method was the introduction by Levene and Van Slyke[52] of the use of barium hydroxide and barium oxide in the liberation of the esters from their hydrochlorides and, later, the use by Foreman[53] of chloroform as solvent for the hydrochlorides, and of anhydrous barium hydroxide for the neutralization. In 1918, Dakin[54] devised the butyl alcohol extraction method as a means to prepare the mixture of mono-amino acids in dry form, suitable for esterification, and applied it, in 1920,[55] to the analysis of gelatin and, in 1923,[56] to the analysis of zein. In this work, the ester method probably reached its highest point of development. A part of the table from this last paper is reproduced in TABLE 11, to illustrate not only the improvements that had been effected over earlier results obtained in Osborne's laboratory, but also to furnish an example of the practice of compiling the highest results of various workers into a table that was supposed to represent the closest approach that had been made to the true composition of the protein. This is a practice that subsequently became common, its chief defense being that only in this way can a complete list of decomposition products be assembled. The more thoughtful are aware of its deficiencies and dangers.

TABLE 11

COMPOSITION OF ZEIN
(Dakin, 1923)

Amino acid	Dakin (1920)	Osborne and Liddle[1] (1910)	Osborne and Clapp[2] (1907)	Highest values
	per cent	*per cent*	*per cent*	*per cent*
Glycine	—	0	0	0
Alanine	3.8	9.8	2.2	9.8
Valine	0	1.9	0.3	1.9
Leucine	25.0	19.8	18.6	25.0
Isoleucine	0	—	—	—
Phenylalanine	7.6	6.6	4.9	7.6
Tyrosine	5.2	3.6	3.6	5.2
Serine	—	1.0	0.6	1.0
Proline	8.9	9.0	6.5	9.0
Oxyproline	0	—	—	—
Aspartic acid	1.8	1.7	1.4	1.8
Glutamic acid	31.3	26.2	18.3	31.3
β-Hydroxyglutamic acid	2.5	—	—	2.5
Tryptophane	0	0	0	0
Arginine	—	1.5	1.2	1.8
Lysine	—	0	0	0
Histidine	—	0.8	0.4	0.8
Ammonia	—	3.6	3.6	3.6
Total		85.5	61.6	101.3

[1] Osborne, T. B., & L. M. Liddle. Am. J. Physiol. 26: 295. 1910.
[2] Osborne, T. B., & S. H. Clapp. Am. J. Physiol. 20: 477. 1907.

THE DETERMINATION OF SINGLE AMINO ACIDS

The observation that proteins yield colored products when treated with certain reagents is very old. Scheele,[57] in 1780, recorded the formation from egg white and from casein that had been treated with acid of what he called "liver of sulfur,"* which was recognized since it darkened silver or lead acetate. He thus detected the presence of sulfur in these proteins by means of the lead-blackening reaction. The earliest workers with nitric acid must have become familiar with the yellow stain this reagent leaves upon the skin, and the general nature of the reaction was mentioned by Welter,[59] in 1799. The term, "xanthoproteic," as applied to this color reaction, was used at least as early as 1836 by Mulder.[60] Tiedemann and Gmelin,[61] in 1826, described a color reaction of enzymatic protein digests with chlorine water, that later[62] became known as the tryptophane reaction. Millon,[63] in 1849, described the reaction with mercuric nitrate still known by his name, the relationship to tyrosine being pointed out by Kühne, in 1868.[64]

---

* A footnote in the German translation of Scheele's works,[58] possibly added by Hermbstädt, the editor, points out that other animal substances such as hair, wool, and silk, after having been dissolved in alkali and treated with acetic acid, likewise give off an "hepatic" odor and that this demonstrates the presence of sulfur in them.

Although these and many other color reactions, some of which had been recognized as being evidence of the presence of specific amino acids or "groups," were in common use, the notion of employing a color reaction as a means for the analytical determination of an amino acid has been a comparatively recent development. Its broad application in the field of protein chemistry is primarily due to Folin. The first comprehensive attempt to determine a single amino acid in a series of proteins by a colorimetric method was that of Folin and Denis[65] who, in 1912, described a method for tyrosine that depends upon the reduction of their phenol reagent with the production of a blue color. The most striking outcome of this work was the observation that the proteins examined all yielded far more tyrosine, frequently twice as much, than had been obtained from them by the earlier isolation procedure. It is beyond the scope of the present discussion to follow the vicissitudes of this pioneer attempt. Suffice it to say that, ten years later, Folin and Looney[66] were able successfully to answer the many critics of the original publication and to add valuable colorimetric methods for the determination of tryptophane and cystine.

The point upon which I wish to dwell in connection with these methods is the novelty of Folin's approach to the problem. Previously, those who had concerned themselves with protein analysis had concentrated almost entirely upon the various groups of amino acids. Kossel and Kutscher worked out the method for the three basic amino acids, which were always determined together. Fischer had dealt with the mono-amino acids, Foreman,[67] with the dicarboxylic amino acids. With the exception of the Hausmann method for basic nitrogen, the methods that had been developed required the use of large quantities of protein, several hundreds of grams for the Fischer method or for the gravimetric determination of tyrosine or cystine; at least 50 gm. for the bases or for the dicarboxylic amino acids. Thus, the range of application of these methods was greatly circumscribed. They could be used only with proteins that could be obtained in large quantities. The sensitive colorimetric methods of Folin permitted determinations to be made upon as little as 1 gm. of material, and, with care, the size of the sample could be diminished still more. Hence, the scope of amino acid analysis was at once greatly enlarged, since preparations of which only small quantities were available could be analyzed. An opportunity was thus presented which has analogies to that which opened before Osborne, when the Hausmann method first appeared. This time, however, the methods were specific ones rather than general.

Another vastly important advance was in the matter of the time required for an analysis. A determination of the basic amino acids, by the Kossel method, required about 4 weeks of hard work, while an ester distillation analysis was a still more time-consuming operation. For both, training and skill of the highest order were necessary, and only a single set of figures was obtained in either case. The colorimetric methods, on the other hand, required only a few hours, aside from the time needed for the hydrolysis, and the manipulations were such that a technical assistant could soon be trained to obtain results as accurate as any. Thus, replication of the analyses became possible, so that average values, together with an estimate of precision, could be secured.

These were fundamental advances in the technique of amino acid analysis, and they were at once appreciated. During the past 25 years, scores of amino acid methods with various modifications and improvements have been developed. They employ all sorts of quantitative techniques: measurements of color values and of optical properties other than color, titrations, weighings on the micro-, semi-micro-, and macro-scale, measurements of electrical properties and of gas volumes, and, most recently, measurements of growth stimulation of specific organisms. The theme that is common to all of these methods is the use of small amounts of protein, and the procedures are all relatively rapid and capable of many repetitions. Folin must be conceded much of the credit for the revolution in point of view that has rendered this development possible.

THE PURPOSE OF AMINO ACID ANALYSIS OF PROTEINS

During the year 1944, a single American journal, the *Journal of Biological Chemistry,* published no less than 21 papers in which new methods, or modifications of previous methods, to determine amino acids were described. This is just over 5 per cent of the 387 papers in the five volumes that appeared, and takes no account of other papers in which such procedures were applied to various problems. The methods dealt altogether with thirteen different amino acids; that is, with about one-half of the substances that are of importance in protein analysis. This is an extraordinary showing, especially if it is remembered that these researches were carried out during the year that this country was exerting its maximum effort in the war. What is the benefit to be derived from such a concentration of study in a field of research that, two decades ago, brought forth hardly a single analogous paper in the same journal? The change is only in part due to an increase in interest in protein chemistry. Many papers on proteins were published in 1924.

It may be helpful to consider, in an attempt to account for this phenomenon, some of the uses that are made of the information derived from the analysis of proteins.

Characterization

The earliest, and still one of the more important applications of amino acid data, is in the characterization of protein preparations. Such information soon became a necessary part of the demonstration that preparations from different sources were different in composition. Later, after this had become largely a matter of assumption, rather than one that required proof, analytical methods were employed in following fractionation experiments, such as those of Haugaard and Johnson,[68] on the components of the gliadin system, in order to demonstrate differences in the successive fractions. Recent examples of papers in the earlier tradition are the demonstrations that serum albumins from different species differ, for example, in tryptophane content,[69] and that the hemoglobin of the bovine fetus differs from that of the adult with respect to histidine.[70]

Composition

It has been almost a matter of faith, for many years, that proteins yield nothing in addition to amino acids after complete hydrolysis. A reservation is made in the case of those proteins which possess a definite prosthetic group, as in hemoglobin, or which, after elaborate purification, have been found to yield carbohydrates or their derivatives in significant proportions. The group of nucleoproteins has become far better understood in recent years, as is also the case with the lipoproteins, although we still do not fully grasp the relation between what may be called the protein part and the rest of the molecule of these more complex substances. Indeed, the very use of the term, molecule, in this connection, may not be regarded as appropriate in all cases. However, for the so-called simple proteins and for the protein moiety of the more complex proteins, the assumption mentioned above is commonly made.

It is well to recall that this is still an assumption. Amino acid analysis is now on the point of attaining its goal of being able to account for all of the carbon, hydrogen, nitrogen, and sulfur of the simple proteins. If there are other amino acids still awaiting discovery, we can be assured that, in at least one protein, as will be shown in a later paper in this conference, they can form only a very small proportion of the molecule. Furthermore, no evidence has been accumulated that suggests the presence of components of other kinds in more than very small proportions. Nevertheless, there are several unsolved problems, such as the origin of the small amount of carbon dioxide produced during acid hydrolysis; and no clear account can even yet be given of the formation of humin.

Stoichiometry

That there is a stoichiometry of at least some of the components of certain proteins is implicit in the early data of Zinoffsky and of Jaquet on the hemoglobins of the horse, dog, and fowl. The iron and sulfur values obtained by these investigators can readily be shown to be in simple integral atomic relationship to each other, as was pointed out by Osborne, in 1902.[71] Osborne studied the total sulfur and the so-called loosely-bound sulfur (sulfur removed as sulfide when the protein is heated with alkali and lead salts) of a series of proteins, and noted a number in which these quantities bore simple integral atomic ratios to each other. Nevertheless, there were many exceptional cases, and few investigators appear to have sought for stoichiometric relationships among the amino acid decomposition products of proteins, until fairly recent years. The attempt by Cohn, Hendry, and Prentiss, in 1925,[72] to compute probable values for the minimal molecular weight of a number of proteins was the first comprehensive application of the stoichiometric relationships among the amino acids. Bergmann and Niemann,[73] eleven years later, developed their well-known periodicity hypothesis of the arrangement of the amino acids in the protein, although this speculation is now held by many to have been somewhat premature. Nevertheless, their views have exerted a marked effect in stimulating analytical research, and the recent discussions of them by Astbury[74] and by Chibnall[75] indicate that further developments are to be anticipated, when the mathematical relationships that are clearly present in many cases are more fully understood. Meanwhile, stoichiometry plays an important role in calculating probable values for the minimal molecular weight[76] and assisting in the selection of the most probable value from among a series of different results in those cases where the amino acid is present in small proportion. For, if our concept of a definite protein molecule is sound, simple integral molecular relationships must be discoverable.

Interpretation of the Chemical Properties of Proteins

The explanation of the many characteristic color reactions of proteins, in terms of the properties of specific amino acid components, has already been mentioned. Such explanations depend upon merely qualitative information regarding the composition of the proteins, but it was the study of color reactions that led to the early conclusions that gelatin is deficient in both tyrosine and tryptophane, and that zein is deficient in tryptophane,[77] facts that were later of importance in accounting for the poor nutritive properties of these proteins.

Innumerable applications of the quantitative data of protein analysis can be pointed out. As a fairly early example may be mentioned the explanation given by Osborne, Leavenworth, and Brautlecht[28] of the quantities of ammonia liberated from proteins respectively by acid and by alkaline hydrolysis, in terms of amides of the dicarboxylic amino acids, in the one case, and in terms of these, together with arginine, in the other case. Choice of suitable illustrations from recent papers is difficult, merely because of the wealth of material available. If a selection is made purely at random, the work of Li[78] on the iodination of serum albumin and pepsin might be mentioned, in which an attempt is made to interpret the quantity of iodine that can be introduced in terms of the proportion of tyrosine radicals in the molecule and the availability of these for reaction, according to whether the protein is dissolved in water or in concentrated urea solution.

There is a large and rapidly growing literature upon such subjects as the reactions of proteins with formaldehyde or quinone, with alkylating and acylating reagents, with esterifying reagents and with general or specific reagents of still other types. In some cases, interpretations can be more or less definitely assigned; in others, the explanation is uncertain or is still to be secured. In all, however, a quantitative knowledge of the reactive groups in the protein molecule, that is, of the amino acid composition, is essential to an understanding of the results.

Rational Explanation of the Physical Properties of the Proteins

That proteins are substances that can form salt-like compounds with both acids and bases, was clear both to Kossel and to Osborne, at the turn of the century, although few of their contemporaries appreciated the meaning of the observations. During the first two decades of the century, the physical-chemical relationships between proteins and acids or bases were developed by Bugarszky and Liebermann,[79] by Hardy,[80] by Michaelis,[81] by Sørensen,[82] and by Loeb.[83] In 1925, the close correlation between the acid-binding capacity and the basic amino acid content, on the one hand, and that between the base-binding capacity and the non-amide bound dicarboxylic amino acid content, on the other hand, were pointed out by Cohn.[84] The validity of such correlations is, of course, entirely dependent upon the accuracy of the physical measurements, which seems moderately satisfactory in some cases, and that of the available analytical results, which, even today, leaves much to be desired. Nevertheless, the success that has attended these efforts indicates the soundness of the theory involved. Many other correlations between physical properties and amino acid composition have likewise received attention. For a discussion of these, the original literature or Cohn and Edsall's monograph[85] should be consulted. Suffice it to say here that such fundamental properties as solubility, mobility, and interaction are now being discussed in terms of the number and distribution of polar and non-polar groups upon the protein molecule; that is, in terms of its amino acid composition and the arrangement of the amino acid radicals upon or within it. Thus, analytical chemistry is the foundation of the present-day development of the physical theory of proteins, and will be the basis of its future extension.

The Structure of the Protein Molecule

In order to arrive at a formulation that expresses the most probable structure of an organic molecule, a definite series of logical steps must be taken. Experience shows that the more complex the molecule, the more difficult it is to obtain evidence that the one and only possible solution has been obtained at each stage in the argument. In general, one starts with an ultimate analysis and an estimation of the molecular weight. The characteristic groups or radicals are then identified qualitatively, and are subsequently quantitatively determined. At this point, it is frequently possible to set up a provisional formulation which may suggest the structure of the fundamental skeleton of the molecule, or may lead to experimental attempts to recognize its nature by decomposition reactions or otherwise. This skeleton having been established, various formulae can be written, among which there are usually a few that remain credible in the light of the properties of the substance. Selection of the one structure that most fully accounts for these properties is, in general, only possible after the hypothetical molecule has been reproduced by synthesis.

It may be useful to consider how far we have progressed, in the case of the proteins, in this inescapable, logical process. As a rule, the greater the molecular weight, the less help is given by the ultimate analysis in the decision between possible alternatives. Even the earliest workers, such as Mulder and Liebig, who attempted to write empirical formulae for the proteins, became impressed with the uncertainties involved. On the other hand, we are today in possession of reasonably accurate estimates of the molecular weight of many typical proteins. The qualitative nature of by far the greater number, and possibly all, of the radicals is also known and there is a considerable body of information concerning the relative proportions of these radicals present in certain of the proteins. It is the next step, the formulation of the fundamental skeleton of the protein molecule, that has so far eluded us, although many hypotheses have been suggested. Of these, the polypeptide hypothesis of Hofmeister and Fischer still occupies the center of the stage, for reasons that are familiar to all. To it has been added the view expressed by Sørensen,[86] in 1930, that the "long peptide chains doubtless will be flexible and apt to roll themselves up by reason of the cohesion forces by which the various parts of the polypeptide chain attract each other." Specific interpretations of this tendency to "roll up" have been assigned by Astbury,[75] by Mirsky and Pauling,[87] by Wrinch,[88] and by others, but hitherto, none of these more detailed views has received general acceptance, nor has any of them been conclusively demonstrated to represent the actual facts, although all are stimulating.

It is clear, then, that we are still at a stage in the development of the theory of protein structure where the data furnished by the analytical chemist are of vital importance. Speculation in this field is one of the most attractive of occupations, and many have yielded to its allure. Nevertheless, it is well to bear in mind that speculation on too narrow a basis of facts should not be regarded as theory.

Metabolism

The discussion of the points concerning which the protein analytical chemist can render service to his colleagues has been restricted, thus far, to those which deal with pure chemistry. There are many applications of the methods of amino acid analysis to biochemistry and physiology. The metabolism of the amino acids in the animal body, as well as in the plant, can be adequately studied only when specific analytical methods for the amino acids themselves and for their products of metabolism have been developed. The increase in interest in this field in recent years is a corollary to the fact that many of these analytical problems have been more or less satisfactorily solved. For example, it was as recently as 1913 that Abderhalden[89] provided the first evidence obtained by isolation that amino acids actually do occur in the blood; at about the same time, Levene and Van Slyke[90] demonstrated their presence in the urine; and Van Slyke and his associates, their presence in the tissues.[91] These observations were largely dependent upon the development by Van Slyke of the convenient nitrous acid method to determine amino acids, and the application of this method in its various modifications has been one of the most important steps in extending our knowledge of amino acid metabolism in both the animal and the plant. That there is still much to learn in this field, is illustrated by the recent observation of Hamilton, in Van Slyke's laboratory,[92] that glutamine is one of the most important amino acid components of blood.

Nutrition

That proteins which differ in their amino acid content as widely as do those of wheat and those of the legumes must also differ in their nutritive effect in the animal, was pointed out by Ritthausen, as early as 1872 (p. 234[9]). This possibility was repeatedly mentioned by Osborne,[27, 93] in early papers dealing with the composition of proteins, and was clearly established by the brilliant nutrition studies of Osborne and Mendel, in the second decade of the century. The doctrine of the indispensability of certain of the amino acids was enunciated by them in 1914.[94] Subsequent research has served to demonstrate the wide, if not universal, application of this principle, and has greatly enlarged the list of amino acids that are required by various species. The continuing investigations of Rose and of Holt, and their respective associates, have broadened our knowledge of details, and have led to the view that protein nutrition is largely, if not entirely, a matter of the supply of the appropriate assortment of amino acids in the most suitable proportions. Thus, the student of nutrition is today concerned as never before with the amino acid composition of food materials, and has turned to the analytical chemist with innumerable and challenging problems. It is, and will increasingly be, the obligation of the amino acid analyst to provide accurate and comprehensive answers.

I should like to close this discussion with a few comments upon the nature of the data that are needed at the present time. It has been shown, I think, that the protein analyst inherits a great tradition. The background of this phase of scientific investigation is illuminated by some of the most eminent names in the history of chemistry and physiology. There is, thus, an obligation to maintain the standards that have long been established in this field. It is also evident that the data frequently have importance far beyond that envisioned by the individual investigator himself. For example, let us consider the case of one who has prepared a concentrate of a new protein hormone and has obtained a few rough estimates of some of its amino acid components, largely as a matter of characterization and as an additional demonstration of its protein nature. The day may come when someone will establish that the mechanism of the specific activity of this hormone involves the presence of certain groups of amino acids in a definite arrangement. He will, accordingly, turn to the literature on its composition. He has two choices: He either has to repeat the analysis himself with due attention to the necessary accuracy, or he must accept the earlier data and take the chance that they will not mislead him into a false interpretation of his own results. This is a purely hypothetical case, but there are myriads of cases in which physical-chemical or physiological principles could be clearly established by examples taken from the literature, if the analytical data were sufficiently reliable and enough were available. Tables of amino acid composition all too frequently contain widely different alternative values for the same amino acid. Such instances imply a degree of uncertainty, with respect to the accuracy of the methods, that renders the information of dubious value. Attempts are, consequently, made to select what are sometimes called the "best" values, but, even with the most careful study of the original papers, it is seldom possible to be assured that the selection has been wise.

Accordingly, the time is ripe for a consideration of what should be expected of data for the amino acid composition of a protein. In the first place, the preparations of protein subjected to analysis must be material of assured purity, as evidenced by correspondence in properties with acceptable standards. Thus, the nitrogen content and the moisture and ash content should invariably be stated, for these are the minimal requirements for such assurance. The origin of the material is also a matter of interest, since, unfortunately, there are few commercial protein preparations sufficiently well standardized to be of real value for these purposes. Those who would modify a former method of analysis or develop a new one then have two definite obligations. They must establish the accuracy of their procedure so that the results are reliable within known and stated limits. This can be done by the recovery of known quantities of the amino acid by itself and from mixtures with other amino acids or protein hydrolysates, or by the analysis of proteins of which the composition is already known. Secondly, they must establish the precision of the method by repetition under all suitable conditions. Obviously, full data must be presented.

Analogous considerations apply to those who employ amino acid methods in studies of metabolism. In addition, it must be established beyond doubt, in this more difficult case, that there is no interference from other components of the material subjected to analysis. If this is unavoidable, steps must be taken to eliminate, or correct for, such interference.

Insistence upon these wholly elementary principles may seem scarcely necessary at this stage, nor would they be dwelt upon, were it not possible to point to dozens of papers in the recent literature where one or more of them has been violated. Further, the very multiplicity of methods for some of the amino acids is evidence for the dissatisfaction that has been felt with many of the procedures. The standards of work in this field are, unfortunately, by no means invariably of the highest, and I believe that it is one of the purposes of the present conference to try to raise these standards to a more acceptable level. If this alone is accomplished, the conference will have been fully justified.

BIBLIOGRAPHY

1. **Braconnot, H.**
   1820. Ann. chim. phys. **13** (2): 113.
2. **Boussingault, J. B.**
   1836. Ann. chim. phys. **63** (2): 225.
3. **Mulder, G. J.**
   1836. Ann. Chem. **20**: 202.
4. **Vickery, H. B., & T. B. Osborne**
   1928. Physiol. Rev. **8**: 393.
   Vickery, H. B.
   1942. Liebig and the chemistry of proteins. Liebig and After Liebig. A.A.A.S. Pub. **16**: 19. Washington, D.C.
5. **Liebig, J.**
   1841. Ann. Chem. **39**: 129.
   1846. Ibid. **57**: 131.
6. **Scherer, J.**
   1841. Ann. Chem. **40**: 1.
7. **Liebig, J.**
   1842. Animal Chemistry, or Organic Chemistry in Its Applications to Physiology and Pathology. Edited by William Gregory. London.
8. **Ritthausen, H.**
   1862. J. prakt. Chem. **85**: 193.
9. **Ritthausen, H.**
   1872. Die Eiweisskörper der Getreidearten, Hülsenfrüchte und Ölsamen. Bonn: 231, 234.
10. **Ritthausen, H.**
    1866. J. prakt. Chem. **99**: 6, 454.

11. **Ritthausen, H.**
1869. J. prakt. Chem. **106:** 445.
1869. Ibid. **107:** 218.
12. **Ritthausen, H., & U. Kreusler**
1871. J. prakt. Chem. **3:** 314.
13. **Liebig, J.**
1846. Ann. Chem. **57:** 127.
1847. Ibid. **62:** 257.
14. **Bopp, F.**
1849. Ann. Chem. **69:** 16.
15. **Hinterberger, F.**
1849. Ann. Chem. **71:** 70.
16. **Guckelberger, G.**
1848. Ann. Chem. **64:** 39.
17. **Schulze, E.**
1880. Landw. Jahrb. **9:** 689.
18. **Hlasiwetz, H., & J. Habermann**
1873. J. prakt. Chem. **7:** 397.
1873. Ann. Chem. **169:** 150.
19. **Osborne, T. B., & G. F. Campbell**
1896. J. Am. Chem. Soc. **18:** 609.
20. **Weyl, T.**
1877. Z. physiol. Chem. **1:** 72.
21. **Drechsel, E.**
1889. J. prakt. Chem. **39:** 425.
22. **Hedin, S. G.**
1895. Z. physiol. Chem. **20:** 186.
1895–1896. Ibid. **21:** 155.
23. **Hedin, S. G.**
1896. Z. physiol. Chem. **22:** 191.
24. **Kossel, A.**
1896. Z. physiol. Chem. **22:** 176.
25. **Kossel, A., & F. Kutscher.**
1900. Z. physiol. Chem. **31:** 165.
26. **Hausmann, W.**
1899. Z. physiol. Chem. **27:** 95.
27. **Osborne, T. B., & I. F. Harris.**
1903. J. Am. Chem. Soc. **25:** 323.
28. **Osborne, T. B., C. S. Leavenworth, & C. A. Brautlecht**
1908. Am. J. Physiol. **23:** 180.
29. **Wells, H. G., & T. B. Osborne**
1911. J. Infect. Dis. **8:** 66.
1913. Ibid. **12:** 341.
1914. Ibid. **14:** 377.
Lake, G. C., T. B. Osborne, & H. G. Wells
1914. J. Infect. Dis. **14:** 364.
30. **Reichert, K. E.**
1849. Arch. Anat. Physiol. u. wiss. Med. **1849:** 198.
31. **Funke, O.**
1851. Z. rat. Med. **1:** 185.
1852. Ibid. **2:** 198, 288.
32. **Hoppe-Seyler, F.**
1862. Arch. path. Anat. u. Physiol. u. klin. Med. **23:** 446.
1864. Ibid. **29:** 223.
1867. Med. chem. Untersuch. **1867:** 167.

33. **Zinoffsky, O.**
1886. Z. physiol. Chem. **10:** 16.
34. **Jaquet, A.**
1888. Z. physiol. Chem. **12:** 285.
1890. Ibid. **14:** 289.
35. **Gamgee, A.**
1898. Textbook of Physiology. Edinburgh & London.
36. **Kossel, A.**
1901. Ber. chem. Ges. **34:** 3214.
37. **Fischer, E.**
1901. Z. physiol. Chem. **33:** 151.
38. **Hopkins, F. G., & S. W. Cole**
1902. J. Physiol. **27:** 418.
39. **Block, R. J., & D. Bolling**
1945. The Amino Acid Composition of Proteins and Foods. Springfield, Ill.
40. **Fischer, E., & E. Fourneau**
1901. Ber. chem. Ges. **34:** 2868.
41. **Schaal, E.**
1871. Ann. Chem. **157:** 24.
42. **Schiff, H.**
1897. Ber. chem. Ges. **30:** 2449.
1899. Ann. Chem. **310:** 301.
43. **Curtius, T., & F. Goebel**
1888. J. prakt. Chem. **37:** 150.
Curtius, T., & H. Schulz
1890. Ber. chem. Ges. **23:** 3041.
44. **Hofmeister, F.**
1902. Ergebn. Physiol. **1:** 759.
45. **Fischer, E.**
1907. Z. physiol. Chem. **53:** 126.
46. **Kossel, A., & H. D. Dakin**
1904. Z. physiol. Chem. **41:** 407.
47. **Taylor, A. E.**
1908–1909. J. Biol. Chem. **5:** 389.
48. **Kossel, A.**
1913. Z. physiol. Chem. **88:** 163.
49. **Osborne, T. B., & S. H. Clapp**
1906. Am. J. Physiol. **17:** 231.
50. **Kossel, A., & A. J. Patten**
1903. Z. physiol. Chem. **38:** 39.
51. **Osborne, T. B., & D. B. Jones**
1910. Am. J. Physiol. **26:** 305.
52. **Levene, P. A., & D. D. Van Slyke**
1908. Biochem. Z. **13:** 440.
1909. J. Biol. Chem. **6:** 419.
53. **Foreman, F. W.**
1919. Biochem. J. **13:** 378.
54. **Dakin, H. D.**
1918. Biochem. J. **12:** 290.
55. **Dakin, H. D.**
1920. J. Biol. Chem. **44:** 499.
56. **Dakin, H. D.**
1923. Z. physiol. Chem. **130:** 159.
57. **Scheele, C. W.**
1786. The Chemical Essays of Charles William Scheele: 265. Translated by T. Beddoes. Edinburgh.

58. **Scheele, C. W.**
1793. Sämmtliche physische und chemische Werke: 249. Edited by S. F. Hermbstädt. Berlin.

59. **Welter**
1799. Ann. chim. **29**: 301.

60. **Mulder, G. J.**
1836. Ann. Physik u. Chem. **37**: 594.

61. **Tiedemann, F., & L. Gmelin**
1831. Die Verdauung nach Versuchen. Second Edition. Heidelberg and Leipzig.

62. **Neumeister, R.**
1890. Z. Biol. **26**: 324.

63. **Millon, E.**
1849. Compt.-rend. Acad. **28**: 40.
1850. Ibid. **29**: 507.

64. **Kühne, W.**
1868. Lehrbuch der physiologischen Chemie: 110. Leipzig.

65. **Folin, O., & W. Denis**
1912. J. Biol. Chem. **12**: 245.

66. **Folin, O., & J. M. Looney**
1922. J. Biol. Chem. **51**: 421.

67. **Foreman, F. W.**
1914. Biochem. J. **8**: 463.

68. **Haugaard, G., & A. H. Johnson**
1930. Compt.-rend. Lab. Carlsberg **18** (2).

69. **Brand, E., B. Kassell, & L. J. Saidel**
1944. J. Clin. Invest. **23**: 437.

70. **Vickery, H. B.**
1944. J. Biol. Chem. **156**: 283.

71. **Osborne, T. B.**
1902. J. Am. Chem. Soc. **24**: 140.

72. **Cohn, E. J., J. L. Hendry, & A. M. Prentiss**
1925. J. Biol. Chem. **63**: 721.

73. **Bergmann, M., & C. Niemann**
1936. J. Biol. Chem. **115**: 77.
1937. Ibid. **118**: 301.
1938. Ibid. **122**: 577.

74. **Astbury, W. T.**
1943. Advances in Enzymology. **3**: 63. New York.

75. **Chibnall, A. C.**
1942. Proc. Roy. Soc. London (B) **131**: 136.

76. **Brand, E., & B. Kassell**
1942. J. Biol. Chem. **145**: 365.

77. **Osborne, T. B., & I. F. Harris**
1903. J. Am. Chem. Soc. **25**: 853.

78. **Li, C. H.**
1945. J. Am. Chem. Soc. **67**: 1065.

79. **Bugarszky, S., & L. Liebermann**
1898. Arch. ges. Physiol. **72**: 51.

80. **Hardy, W. B.**
1905. J. Physiol. **33**: 251.

81. **Michaelis, L.**
1914. Die Wasserstoffionenkonzentration. Berlin.

82. **Sørensen, S. P. L.**
1915–1917. Compt.-rend. Lab. Carlsberg **12**.

83. **Loeb, J.**
1922. Proteins and the Theory of Colloidal Behavior. New York.

84. **Cohn, E. J.**
1925. Physiol. Rev. **5**: 349.

85. **Cohn, E. J., & J. T. Edsall**
1943. Proteins, Amino Acids and Peptides. New York.

86. **Sørensen, S. P. L.**
1930. Compt.-rend. Lab. Carlsberg **18** (5).

87. **Mirsky, A. E., & L. Pauling**
1936. Proc. Nat. Acad. **22**: 439.

88. **Wrinch, D. M.**
1937. Proc. Roy. Soc. London (A) **160**: 59.
1940. Phil. Mag. **30**: 64.
Wrinch, D. M., & I. Langmuir
1938. J. Am. Chem. Soc. **60**: 2247.

89. **Abderhalden, E.**
1913. Z. physiol. Chem. **88**: 478.

90. **Levene, P. A., & D. D. Van Slyke**
1912. J. Biol. Chem. **12**: 301.

91. **Van Slyke, D. D., & G. M. Meyer**
1913–1914. J. Biol. Chem. **16**: 197.

92. **Hamilton, P. B.**
1942. J. Biol. Chem. **145**: 711.
1945. Ibid. **158**: 397.

93. **Osborne, T. B., & R. D. Gilbert**
1906. Am. J. Physiol. **15**: 333.

94. **Osborne, T. B., & L. B. Mendel**
1914. J. Biol. Chem. **17**: 325.

THE USE OF SPECIFIC PRECIPITANTS IN THE AMINO ACID ANALYSIS OF PROTEINS

By William H. Stein and Stanford Moore

The Rockefeller Institute for Medical Research, New York, N.Y.

The immediate goal which protein chemists hope to attain with the aid of amino acid analysis is the ability to express the empirical formulae of numerous proteins in terms of their constituent amino acids, in much the same manner as the empirical formulae of simpler compounds are expressed in terms of their constituent elements. The protein chemist, as a consequence, should be able to analyze a protein for its constituent amino acids with the accuracy and dispatch attending the elementary analysis of simple compounds. Accuracy primarily, and secondarily dispatch, therefore, will be the criteria by which the methods to be discussed in this communication will be judged.

In this discussion, a "specific precipitant" will be defined as any reagent which, when added to a protein hydrolysate, will unite with, and cause to precipitate, one or a small group of amino acids by the formation of insoluble salts or molecular complexes. Reagents which combine chemically with amino acids through covalent linkages to form insoluble derivatives will not be considered.

The use of precipitants for the exploration of the amino acid composition of protein hydrolysates has a long history. This history may be divided roughly into its qualitative and quantitative aspects. Our extensive knowledge of the distribution of amino acids in nature is a result largely of the qualitative or semi-quantitative use of specific precipitants. This phase of the subject has been dealt with elsewhere,[1, 2, 3] and hence will not be discussed here. We shall be concerned, rather, with the quantitative aspects of the subject. In the past, many specific precipitants have been employed as quantitative tools for the determination of specific amino acids, or small groups of amino acids, in protein hydrolysates. The information thus obtained has been of the greatest importance, and is the foundation upon which the newer quantitative work must build. The question may be raised, however, whether quantitative isolation procedures offer promise of future rapid advances in the amino acid analysis of proteins. In the course of this discussion, an attempt will be made to answer this question. In order to do this, an evaluation of methods is necessary. The methods selected are those which appear to be the most accurate; which have come to be more or less widely employed for the determination of single, or small groups of amino acids; and which depend for their success, or lack of it, solely on selective precipitation techniques.

In order to start the discussion of "specific precipitants," the aromatic sulfonic acids have been chosen. This departure from chronology has been undertaken, not with any exaggerated belief in the importance of these reagents, but rather because the investigations to be described bring out clearly certain fundamental principles regarding the behavior of precipitants in general.

The use of aromatic sulfonic acids as precipitants for amino acids dates from the discovery by Kossel and Gross, in 1924,[4] that flavianic acid (2,4-dinitro-1-naphthol-7-sulfonic acid) forms insoluble salts with the basic amino acids. Prior to this, Suida[5] described salts of the basic amino acids with two sulfonated dyes, Orange II and crystal Ponceau. It was Kossel and his colleagues, however, who first employed a sulfonic acid precipitant for the determination of an amino acid in a protein hydrolysate. Subsequently, Zimmermann[6] described salts of rufianic acid (1,4-dihydroxyanthraquinone-2-sulfonic acid) with basic substances and with tyrosine and phenylalanine, and later[7] described salts of purpurin sulfonic acid. Ackermann[8] observed that 2,6-diiodophenol-4-sulfonic acid precipitates the basic amino acids.

The earlier investigators were apparently drawn to the sulfonic acids because of their use as dyes or vital stains, and thought of them primarily as precipitants for the basic amino acids. The aromatic sulfonic acids as a class, however, are so strongly acidic that they may be expected to form salts with all types of amino acids. In the laboratory of the late Dr. Max Bergmann, it was found[9] that naphthalene-β-sulfonic acid forms sparingly soluble salts, not only with the basic amino acids, but with leucine and phenylalanine, as well. This observation prompted the investigation[10, 11] of the solubilities of a considerable number of amino acid salts of sulfonic acids, in the hope that new reagents suitable for the determination of amino acids might be found. The solubilities at 0° of the amino acid salts of several sulfonic acids are given in TABLES 1 and 2.

The procedure employed to obtain these data was an admittedly approximate one, and need not be detailed here. The method permitted a rapid comparative investigation of the behavior of over 100 sulfonic acids, each with 15 to 20 different amino acids. Sulfonic acid reagents which appeared of interest on the basis of these preliminary experiments were investigated in greater detail. In such cases, the salts were prepared, analyzed, and their solubilities accurately determined. Such accurately determined solubilities are underlined in the tables.

The reasons for listing the solubility products in the tables, rather than the solubilities in grams per liter, will emerge in the course of this discussion. For purposes of orientation, it may be noted that, for the average binary salt of an amino acid, a solubility product of 4×10^{-4} represents a solubility of the order of 1 per cent; 4×10^{-2}, 10 per cent; and 4×10^{-6}, 0.1 per cent. Where no value is given in the tables, the solubility product of the salt in question was found to be higher than 4×10^{-2}, or greater than about 10 per cent.

In examining the data in the tables, it should also be pointed out that the arginine, histidine, lysine, and cystine salts are ternary salts, containing two equivalents of sulfonic acid for each equivalent of amino acid. For these salts, therefore, the solubility product constant is given by the expression:

$$K = [\text{sulfonic acid}]^2 \times [\text{amino acid}].$$

For the other amino acid salts, however, which are binary salts, the sulfonic acid concentration enters into the solubility product constant only as the first power:

$$K = [\text{sulfonic acid}] \times [\text{amino acid}].$$

For this reason, a comparison of the solubility product of a ternary salt with that of a binary salt does not provide a direct measure of their relative solubilities. For example, the *l*-arginine and *dl*-phenylalanine salts of 2,5-dibromobenzene sulfonic acid have approximately the same solubility products. A saturated solution of the ternary arginine salt, however, contains about four times as many moles of amino acid per liter as a saturated solution of the binary phenylalanine salt.

Even on most casual inspection of the data given in the tables, certain conclusions are apparent. In the first place, it can be seen that a large number of aromatic sulfonic acids form insoluble salts with amino acids. The generality of this phenomenon suggests that the possibilities of finding useful reagents among the sulfonic acids have by no means been exhausted.

It will also be noted from the tables that the insoluble salts are not limited to those in which the basic amino acids participate. In fact, there is at least one sulfonic acid listed which forms a sparingly soluble salt with any one of the amino acids. Insoluble salts of the basic amino acids, indeed, are not of such frequent occurrence as are those of leucine and phenylalanine. Unfortunately, however, the solubilities of the salts vary in an unpredict-

Table 1

SOLUBILITY PRODUCTS OF AMINO ACID SALTS OF AROMATIC SULFONIC ACIDS

Amino Acid	Solubility Products of Sulfonic Acid Salts										
	p-Chloro-benzene-	p-Bromo-benzene-	p-Iodo-benzene-	*2,5-Di-bromo-benzene-	*2,5-Di-chloro-benzene-	3,4-Di-chloro-benzene-	4-Nitro-tolu-ene-2-	2-Bromo tolu-ene-5-	*2,6-Di-iodo-phenol-4-	2,4-Dinitro-1-naphthol-7-(flavianic)	*5-Nitro-naphtha-lene-1-
l-Alanine											
l-Arginine				2.5×10^{-5}		7.5×10^{-7}			7.8×10^{-6}	1.3×10^{-7}† 2.4×10^{-8}‡	1.6×10^{-6}
l-Aspartic acid											
l-Cysteine	1×10^{-2}	5×10^{-3}	2×10^{-3}					2×10^{-3}			
l-Cystine										1×10^{-6}	
l-Glutamic acid											
Glycine							2.8×10^{-2}			2.5×10^{-3}	5.2×10^{-4}
l-Histidine				2.5×10^{-5}	1×10^{-3}	1.9×10^{-7}		2×10^{-4}	2×10^{-7}	2.2×10^{-5}§ 5.0×10^{-6}¶	1.6×10^{-6}
l-Hydroxyproline									8.5×10^{-3}		8.7×10^{-4}
l-Isoleucine											
l-Leucine	3.3×10^{-2}	1.2×10^{-2}	2.8×10^{-3}	5.7×10^{-3}	1×10^{-2}	2.6×10^{-3}		9.3×10^{-4}	1.7×10^{-3}	2.9×10^{-4}	1.3×10^{-2}
l-Lysine										1.3×10^{-6}	2.5×10^{-5}
dl-Methionine	1×10^{-2}	1×10^{-2}	1.6×10^{-3}		5×10^{-3}	1.6×10^{-3}					
dl-Phenylalanine	1.2×10^{-2}	3.5×10^{-3}	4.2×10^{-4}	3.2×10^{-5}	4.1×10^{-4}	4.8×10^{-4}	2×10^{-3}	1×10^{-3}‖	1.9×10^{-4}	9×10^{-5}‖	6.7×10^{-5}
l-Proline									6.8×10^{-3}		
dl-Serine											
l-Tyrosine	5.1×10^{-3}	1.9×10^{-2}		1.7×10^{-3}	8.2×10^{-3}		2.8×10^{-3}		7.8×10^{-4}	2.8×10^{-6}	
l-Valine											

* The salts of *l*-cysteine, *l*-cystine, and *dl*-methionine were not investigated in these cases.
† Arginine monoflavianate.
‡ Arginine diflavianate.
§ Histidine monoflavianate.
¶ Histidine diflavianate.
‖ *l*-Phenylalanine salts.

TABLE 2

SOLUBILITY PRODUCTS OF AMINO ACID SALTS OF AROMATIC SULFONIC ACIDS

Amino acid	Solubility Products of Sulfonic Acid Salts										
	Azobenzene-p-	p-Hydroxyazobenzene-p′-	p-Hydroxyazobenzene-m′-	2-Hydroxy-5-methylazobenzene-3′-	4-Nitro-4′-methyldiphenylamine-2-	2-Nitro-4′-methyldiphenylamine-4-	4-Nitrodiphenylamine-2-	Anthraquinone-α-	Anthraquinone-β-	5-Nitroanthraquinone-1-	Picryl
l-Alanine	4×10^{-4}	2×10^{-3}						3×10^{-3}	2×10^{-3}	6×10^{-4}	
Ammonia	4×10^{-2}	4×10^{-5}	3×10^{-4}	3×10^{-3}	4×10^{-2}	2×10^{-3}	4×10^{-4}	6×10^{-5}	3×10^{-3}	1×10^{-4}	4×10^{-4}
l-Arginine	2×10^{-8}	2×10^{-8}		4×10^{-10}	Oil	4×10^{-9}	3×10^{-7}	2×10^{-8}	3×10^{-9}	2×10^{-9}	8×10^{-6}
l-Aspartic acid										2×10^{-8}	
l-Cysteine	2×10^{-5}	8×10^{-6}		2×10^{-6}		2×10^{-3}		6×10^{-4}		4×10^{-5}	
l-Cystine	1×10^{-6}	8×10^{-8}		6×10^{-8}	Oil		Oil	8×10^{-7}	3×10^{-8}	8×10^{-9}	
l-Glutamic acid		4×10^{-3}		1×10^{-2}							
Glycine	2×10^{-4}	1×10^{-3}		3×10^{-5}			5×10^{-4}	2×10^{-4}	1×10^{-2}	4×10^{-6}	3×10^{-3}
l-Histidine	3×10^{-8}	2×10^{-9}		5×10^{-9}	2×10^{-8}	2×10^{-8}	3×10^{-7}	3×10^{-9}	3×10^{-8}	8×10^{-9}	2×10^{-5}
l-Hydroxyproline				1×10^{-2}						3×10^{-8}	
l-Isoleucine	4×10^{-4}			1×10^{-2}	8×10^{-6}	1×10^{-4}		3×10^{-4}	2×10^{-4}	3×10^{-3}	
l-Leucine	8×10^{-5}	3×10^{-4}		8×10^{-6}	2×10^{-4}	8×10^{-5}	5×10^{-4}	5×10^{-4}	1×10^{-4}	2×10^{-5}	
l-Lysine	1×10^{-7}	3×10^{-8}		1×10^{-7}	Oil	4×10^{-7}	6×10^{-9}	6×10^{-8}	1×10^{-6}	3×10^{-9}	3×10^{-4}
dl-Methionine	6×10^{-5}	2×10^{-4}		7×10^{-6}	7×10^{-6}	4×10^{-5}		1×10^{-4}	3×10^{-3}	2×10^{-4}	
l-Phenylalanine	1×10^{-5}	6×10^{-6}	4×10^{-5}	1×10^{-6}	5×10^{-4}	2×10^{-4}	3×10^{-5}	1×10^{-3}	3×10^{-5}	6×10^{-5}	
l-Proline		3×10^{-4}		8×10^{-4}	3×10^{-5}					3×10^{-4}	5×10^{-4}
l-Serine		2×10^{-4}						2×10^{-8}		4×10^{-8}	
dl-Threonine										6×10^{-4}	
l-Tyrosine		1×10^{-4}							2×10^{-8}		
l-Valine	1×10^{-4}			4×10^{-8}	1×10^{-5}	8×10^{-5}		1×10^{-3}	4×10^{-3}	1×10^{-8}	

able manner, with the structure of both the amino acid and the sulfonic acid involved. The search for useful reagents remains, therefore, largely on a trial and error basis.

On closer inspection of the quantitative aspects of the data presented in TABLES 1 and 2, it may be predicted that none of the sulfonic acids is capable of selectively and quantitatively precipitating a given amino acid from the complex mixture of closely related compounds presented by the typical protein hydrolysate. Those reagents which form very insoluble salts with one amino acid do so, as well, with one or more other amino acids, thus enhancing the danger of coprecipitation. Those reagents which precipitate only a few amino acids form salts of quite appreciable solubility. When using these reagents, therefore, in order to obtain yield, purity must be sacrificed. In order to obtain purity, yield must suffer. As we shall see, the degree to which these conflicting factors can be reconciled limits and determines the accuracy of all quantitative isolation procedures for the estimation of amino acids.

In assessing the accuracy of quantitative isolation procedures, then, the two factors, yield and purity, must be evaluated. In appreciation of these factors, and in an attempt to increase the yield of a given amino acid salt, arbitrary solubility corrections have often been employed in order to account for the solubility of a salt in a hydrolysate. But the solubility of a salt in a solvent is a function of the nature of the solvent. A protein hydrolysate, with its dipolar ions, its Cl^- or $SO_4^=$, its H^+, and its other constituents, all present in an unknown concentration, is, in fact, a solvent of undefined properties. The application to such an unknown solvent of solubility corrections, determined in water or some other closely defined solvent, is always a potential source of error.

More attention, in general, has been paid to yield than has been paid to purity. In fact, it has often been said of isolation procedures that they at least give minimal figures. This is not always true, however. High results may be obtained if the product isolated is not pure. Ascertainment of the purity of an amino acid or an amino acid salt with a high degree of precision is not always readily accomplished. Elementary analysis for one element and optical rotation do not always suffice. Frequently, a contaminated product will give figures so close to those of the desired amino acid that appreciable quantities of impurities may be overlooked. Moreover, melting points on amino acid salts, particularly when accompanied by decomposition, cannot be regarded as satisfactory criteria of purity. Solubility techniques, depending upon the phase rule, have been of considerable utility in proving the purity of amino acids and their salts.

The task of evaluating the accuracy of methods for amino acid analysis would be rendered far easier, if each method were checked by the analysis of a known mixture of amino acids approaching in complexity the composition of a protein hydrolysate. This important check was employed by Osborne and Jones,[12] in an evaluation of the Fischer ester method. To our knowledge, however, similar check analyses have not been performed with most of the methods to be discussed here. The lack of unequivocal data on known mixtures of amino acids, probably as much as any other single factor, is responsible for much of the confusion and debate to be found in the literature.

In the preceding discussion, examination has been made of the principles involved in the performance of quantitative isolation procedures, using as a model the sulfonic acids. To evaluate some of the individual methods now in use, we return to the terms used earlier: accuracy, primarily, and dispatch. In evaluating accuracy, particular attention will be paid to the yield and purity of the products isolated. Wherever it is justifiable, an attempt will be made to compare the results obtained by different methods of analysis. The results obtained with the isotope dilution method have been selected as a primary standard with which to compare, whenever possible, the values secured with other methods. This choice may appear somewhat arbitrary, but we believe it to be justifiable.

Two methods in which a sulfonic acid plays a paramount role may be discussed first. They are the monoflavianate method of Kossel and Gross[4] for the determination of arginine, and the diflavianate method of Vickery[13] for the same amino acid. The possible sources of error may be fairly clearly defined in these procedures, and it is of illustrative value to consider the methods in some detail.

In the original Kossel and Gross procedure, arginine monoflavianate was precipitated directly from the neutralized hydrolysate, recrystallized and weighed. This method has been used in several laboratories, but, as Vickery has pointed out, no thorough study of the conditions most suitable for the determination has been made. On the basis of his own work, Vickery concluded that the arginine monoflavianate obtained directly from a hydrolysate was frequently contaminated and difficult to purify.

In the diflavianate procedure of Vickery, a large excess of flavianic acid is used, and arginine is precipitated as the diflavianate directly from a hydrolysate sufficiently acid (pH 1 or less) to insure that both acid-binding groups of arginine are largely in the ionic state. The diflavianate is then transformed into the monoflavianate and weighed. Purity of the monoflavianate is assayed by sulfur analysis. Employing this method, Vickery has analyzed a number of proteins and has compared his most recent results with others obtained by the silver method, the Van Slyke nitrogen distribution method, the arginase method, and the monoflavianate method. For the most part, the results obtained with the diflavianate method are higher than those obtained with the silver or monoflavianate methods. Good agreement between the diflavianate and the arginase method was, in general, observed.

Before proceeding further with an evaluation of the diflavianate method, certain facts pertaining to flavianic acid and its amino acid salts should be noted. It can be seen from TABLE 1 that flavianic acid forms both mono- and di-salts with arginine; and, further, that both of these salts have a very low solubility. So far as we know now, flavianic acid is unique in this respect, although detailed information on this point is available for only a limited number of the sulfonic acids listed. The monoflavianate, which forms in neutral or moderately acid solution, has a solubility product of 1.3×10^{-7}. A saturated solution of the pure salt contains at 0° about 60 mg. of arginine per liter. The diflavianate, which forms in strongly acid solution, has a lower solubility product, 2.4×10^{-8}, but a saturated solution of this compound contains at 0° about 320 mg. of arginine per liter. However, the flavianic acid enters into the solubility product equation as the square in the case of the di-salt, whereas it enters only as the first power in the case of the mono-salt. For this reason, the quantity of arginine remaining in solution at equilibrium is much more sensitive to the flavianic acid concentration in the case of the di-salt than it is in the case of the mono-salt. For practical work, therefore, when a considerable excess of flavianic acid is employed, the extent of the precipitation of arginine may be equally great and, in fact, may be virtually quantitative, no matter which compound is isolated. A large excess of flavianic acid, however, would be expected to cause co-precipitation, for it will be noted that the tyrosine, leucine, phenylalanine, histidine, lysine, and cystine salts of flavianic acid are all quite insoluble.

In the light of these facts, the following observations may be made in regard to the diflavianate method. In practice, the method takes advantage of the mass law. A large excess of flavianic acid is used, so that the calculated solubilities given above may not give an accurate impression of the possible losses of arginine. Under such conditions, it can be calculated and demonstrated, as well, that with pure arginine solutions the precipitation as diflavianate involves virtually no loss, if equilibrium conditions are attained. Whether the same holds true with a protein hydrolysate, however, is difficult to say. For example, protein hydrolysates come to equilibrium very slowly, 5 or more hours of continuous and highly efficient agitation sometimes being required. Moreover, even in the presence of excess flavianic acid, the solubility of arginine diflavianate in a hydrolysate of 10 per cent concentration may be much greater than it is in dilute HCl solution under the same conditions. On the other hand, as might be predicted from the data given above, there can be little doubt that the arginine diflavianate thus obtained is not pure. It is primarily to remove the impurities that the diflavianate is next transformed to the monoflavianate. This procedure involves a fundamental change in experimental conditions. Moreover, as was pointed out above, the precipitation of arginine as monoflavianate is virtually complete, without a large excess of flavianic acid being present. The precipitate is thus less likely to be contaminated by other amino acids carried down with the diflavianate precipitate. However, the purity of the product is not proved beyond question. Sulfur analyses, which were performed, may not demonstrate the presence of other amino acid flavianates in the arginine salt, because the sulfur content of most amino acid flavianates is almost the same. Thus, arginine monoflavianate contains 6.56 per cent S; histidine monoflavianate, 6.82 per cent S; histidine diflavianate, 8.17 per cent S; and tyrosine monoflavianate, 6.46 per cent S. It may be noted that Chibnall *et al.*[14] have presented evidence indicating that the arginine monoflavianate thus prepared is not pure, and requires several recrystallizations.

Despite the possible sources of error which have been pointed out, the diflavianate method has probably yielded the most accurate results obtained from any of the quantitative isolation procedures for the determination of amino acids. In this connection, the data given in TABLE 3 are of interest. β-Lactoglobulin and horse hemoglobin have each been analyzed

TABLE 3

ARGININE CONTENT OF HORSE HEMOGLOBIN AND β-LACTOGLOBULIN

Author	Method	Arginine Content of β-Lactoglobulin	Arginine Content of Horse Hemoglobin
		per cent	per cent
Vickery[13]	Diflavianate	2.66	3.59
Foster[15]	Isotope Dilution		3.7
Chibnall *et al.*[14]	Phosphotungstic-Tristram-Block	2.89	3.75*
Cannan *et al.*[16]	Tristram-Block	2.79	
Brand and Kassell[17]	Colorimetric	2.87	

* Determined in Chibnall's laboratory by the diflavianate method of Vickery.

for arginine by several different methods.[13, 14, 15, 16, 17] As can be seen, the agreement is impressive; and, doubtless, the arginine content of these protein hydrolysates is known within very narrow limits. It should be noted, however, that the arginine content of the proteins in question is low. In this range, systematic errors of the order of 5 to 10 per cent, which might be involved in a given method, would be masked, since they would fall within the limits of error of almost any method. Final proof of the general accuracy of all the quantitative isolation procedures mentioned, therefore, should await more data of a similar nature obtained on several proteins of higher arginine content.

As for the dispatch with which results can be obtained, it may be noted that the method is much simpler than the silver method for arginine, although the manipulations are still fairly numerous. Furthermore, relatively large amounts of material, by modern micro standards (2 to 5 gm.), are required for each single analysis.

Leaving temporarily the subject of the sulfonic acids, we turn next to the estimation of arginine and histidine through the use of silver salts. The ability of the basic amino acids to form sparingly soluble silver salts has been recognized since the discovery of these amino acids, late in the last century. Kossel and Kutscher[18] utilized this property in their method for arginine and histidine. Since its original use by Kossel and Kutscher, the silver method has been modified by almost every one who has had extensive experience with it. The modifications have been directed toward increasing the selectivity of the silver precipitation, and toward improving the yield of arginine and histidine isolated from the final fractions.

Probably the most fundamental modifications of the silver method have been proposed by Vickery and his colleagues. Vickery and Leavenworth[19-23] showed that histidine-silver is insoluble at pH 7.0, whereas arginine silver began to be precipitated at a more alkaline pH, thus providing a separation of the two bases. The same workers also introduced flavianic acid into the method for the final isolation of both arginine and histidine. Osborne, Leavenworth, and Brautlecht,[24] and Osborne, Van Slyke, Leavenworth, and Vinograd[25] first precipitated all of the basic amino acids as phosphotungstates, prior to their separation by means of the silver method. As we shall see later, Chibnall returns to this type of procedure in his most recent work. Among other modifications may be listed the use of $HgSO_4$ for the purification of histidine, proposed by Kossel and Patten;[26] the use of nitranilic acid for the final isolation of histidine, by Block;[27] the adaptation of the method to a small scale, by Block;[28] modifications of this latter procedure, by Tristram;[29] and, finally, the introduction of 3,4-dichlorobenzene sulfonic acid for the isolation of histidine, by Vickery.[30]

So far as we are aware, no analyses of complex known mixtures of amino acids have been reported for any of the modifications of the method listed above. There are no detailed data on the solubilities of the silver salts of all amino acids, data of the type given for the sulfonic acids. As a result, it is not possible to predict the nature and extent of the contamination of the silver fractions with the silver salts of other amino acids. For example, the silver salts of aspartic acid and glutamic acid are moderately insoluble, and, in fact, these amino acids have been isolated from the histidine fraction.[23] Cystine or cysteine are almost invariable contaminants of the histidine fraction.[31] The uncertainties already mentioned in regard to the solubilities of salts in hydrolysates, and the doubts concerning empirical solubility corrections, are once more encountered. Moreover, the purity of the final products has, in general, been assayed by sulfur analyses, and is, therefore, subject to the uncertainties already mentioned. Finally, the silver method is a very complicated procedure.

As was mentioned earlier, Vickery has shown that analyses for arginine performed by the diflavianate procedure gave notably higher results, with many proteins, than had been obtained by the silver method. In the case of histidine, the results of Albanese,[32] who used an electrolytic method, of Brand,[33] who used a colorimetric method, and of Dunn,[34] who used a microbiological method, are in general higher than those obtained with the silver method. These comparative data are not extensive, but those who have had most experience with the silver method for histidine have themselves expressed grave doubts as to its accuracy. We must, therefore, await the results obtained by the isotope dilution method before making any final judgment.

We turn next to the use of phosphotungstic acid as a precipitant for the basic amino acids, with particular emphasis on the determination of lysine, and the work of Chibnall *et al.*[14] on the determination of the basic amino acids. In a recent paper, Van Slyke, Hiller, and Dillon[35] have studied the solubilities of the phosphotungstates of the basic amino acids, and of proline, glycine, and tryptophane. The opening paragraphs of their paper sum up the history of phosphotungstic acid so concisely and admirably that they may be quoted in part here:

"Phosphotungstic acid (which will be designated as PTA) has long been used in separations of the amino acids. It was employed by Drechsel,[36] in the discovery of lysine in casein, by Hausmann,[37] by Osborne, Leavenworth, and Brautlecht,[24] and by Van Slyke,[25, 38-41] for quantitative precipitation of the diamino acids, arginine, lysine and histidine, in protein hydrolysates.

"The monoamino acids also form crystalline phosphotungstates. Although many times more soluble than the PTA salts of the diamino acids, some of the monoamino phosphotungstates have been found useful for separations of individuals within the monoamino group.[42]

"The phosphotungstates of the diamino acids, when precipitated under the conditions customarily employed with protein hydrolysates,[37-41] *viz.*, about 1 N HCl or H_2SO_4 and room temperature, have solubilities sufficiently great to necessitate corrections for them in accurate quantitative work.[38-40] The effects of varying temperature, acidity, and of coprecipitation of the different diamino acids together have, however, not heretofore been subjected to a systematic study. It appeared that such a study might increase the accuracy of solubility corrections, and perhaps also lead to discovery of conditions of precipitation which would make the phosphotungstate separation of the diamino acids from the monoamino group more sharp and certain."

On the basis of their work, Van Slyke, Hiller, and Dillon conclude that the most complete separation of the diamino acids from the monoamino acids should be obtainable by precipitation of the diamino acids with PTA in large volumes at low temperatures. They also noted that the completeness of the precipitation is influenced by the HCl concentration, and that an excess of PTA is necessary. The authors give solubility corrections, determined in 0.25 N HCl which should be applied to the determination of the basic amino acids in protein hydrolysates. The influence of amino acids, other than the ones mentioned above, upon the solubilities of the PTA salts was not investigated, nor was the completeness of the precipitation of the basic amino acids from complex known mixtures of amino acids.

It follows from the work of Van Slyke *et al.*, which appeared in 1942, that the values obtained with the aid of PTA prior to that date are subject to uncertainties of a magnitude not readily ascertainable. These uncertainties involve incomplete precipitation of the basic amino acids as PTA salts, and inaccurate solubility corrections applied to account for this fact. These considerations are significant in the case of the determination of lysine by the method of Kossel and Kutscher, in which lysine is precipitated as PTA salt from the filtrates remaining after removal of arginine and histidine from a hydrolysate. The lysine is subsequently isolated and weighed as the picrate.

According to Tristram,[29] lysine picrate has a solubility in water of 3.4 mg. per cc. at 0°, and, as usually obtained, must be recrystallized in order to attain purity. Lysine picrate, therefore, is an appreciably soluble salt, and account must be taken of this fact. Whether solubility corrections are applicable will depend to some extent on how selective the initial PTA precipitation has been, since other amino acids in the lysine fraction will alter the solubility of the picrate. Moreover, picric acid is a notoriously unselective reagent, so that contaminants brought down by the PTA may easily carry through to the final product. The purity of this final product is frequently assayed by its explosion point, material which explodes sharply over 260° being taken as pure. No data are available, however, on the nature and extent of the impurities which must contaminate lysine picrate before it will fail to explode above 260°.

Unfortunately, there are few data of proven accuracy on the lysine content of proteins with which the PTA results can be compared. The results obtained by electrolytic,[32] microbiological,[44] enzymatic,[43, 45] and isotope dilution methods[46] are, for the most part, notably higher than those obtained by isolation. It would appear probable, therefore, that the existing quantitative isolation methods for the determination of lysine in most cases yield figures which are low.

Phosphotungstic acid as a precipitant for the basic amino acids has also been used by Chibnall and his co-workers[14, 47] in their recent work on the determination of arginine, histidine, lysine, aspartic, and glutamic acids. A discussion of the procedure employed for the basic amino acids, however, is not feasible, until the work on the dicarboxylic amino acids has been reviewed, and the principles of the method as a whole have been stated.

The method employed by Chibnall for the estimation of the dicarboxylic acids is a modification of the Foreman lime-ethanol method.[48] The Foreman method, in turn, is based on observations, made much earlier,[1] on the insolubility in alcohol of the calcium and barium salts of aspartic and glutamic acid. Originally, Foreman precipitated the calcium salts of aspartic and glutamic acid from a neutralized hydrolysate by the addition of alcohol. Other investigators, notably, Dakin,[49] and Jones and Moeller,[50] employed barium instead of calcium as a precipitant. In almost all cases, aspartic acid is finally isolated as the copper salt, and glutamic acid as the hydrochloride.

On the basis of their recent exhaustive studies, Chibnall *et al.* have drawn several important conclusions. They concluded, first, that the calcium salts possess a selectivity superior to that of the barium salts. They also found that unexpectedly large amounts of the dicarboxylic acids escape precipitation in the original lime-ethanol treatment. In order to circumvent this difficulty, the filtrate from the first lime-ethanol treatment is reworked. The basic amino acids are precipitated with PTA, and most of the neutral amino acids are removed, either by direct precipitation, or by use of the Town copper salt method.[51] The resulting solution is then subjected to a second lime-ethanol treatment and a second PTA precipitation. The final precipitations are thus carried out from a solution containing only minimal amounts of the monoamino acids. Chibnall *et al.* believe that, by proceeding in this manner, the solubility factors inherent in the lime-ethanol and PTA precipitations are overcome. The PTA precipitates are worked up for the basic amino acids by the Tristram-Block procedure. Chibnall *et al.* claim an accuracy of 2 per cent for the values obtained by this method.

The work of Chibnall *et al.* has greatly simplified discussion of the methods for aspartic and glutamic acids. According to these workers, "these findings show, without any equivocation, that the lime-ethanol method of Foreman, as applied by us and by all other workers in the past, must have given results that were far from quantitative. . . ." Remaining to be discussed, therefore, is only the Chibnall procedure itself. A detailed analysis of this monumental work is scarcely possible here, nor is it necessary. The flow sheet published with the method reveals 29 major steps and numerous minor ones. On the basis of what has been said of the many possible sources of error inherent in quantitative isolation methods, it may be doubted whether a procedure of such complexity can yield results accurate to 2 per cent. Moreover, the theoretical doubts already voiced are borne out by the recent work of Foster[15] with the isotope dilution method, and of Snell[52, 53] with new microbiological methods. The results obtained by Chibnall, Foster, and Snell for the glutamic and aspartic acid content of several proteins are given in TABLE 4. It may be noted that the values of Chibnall for glutamic acid are usually somewhat higher than are those of Foster or of Snell. The latter two workers analyzed only for *l*-glutamic acid, whereas Chibnall claims to have isolated appreciable amounts of *dl*-glutamic acid. Previously, it had been shown,[54] by the isotope dilution technique, that the glutamic acid in hydrolysates of tumor and normal tissue was 99 per cent in the *l*-form. If the same state of affairs holds for all proteins, Chibnall's *dl*-glutamic acid might represent a mixture of *l*-glutamic acid and other amino acids. The question encroaches on many controversial matters which cannot be dealt with here. It might be well, therefore, to reserve final judgment, pending further work with the isotope dilution method. Similar uncertainties are not encountered in the case of aspartic acid. Here, Chibnall's figures are much lower than are those of Foster or of Snell.

As regards the accuracy of the determinations of the basic amino acids by this procedure, unequivocal judgment is difficult. The determination of arginine is, doubtless, the most accurate, and little more need be added to what has already been said on this subject. In the case of histidine and lysine, it may be questioned whether the elaborate technique of Chibnall *et al.* has succeeded in eliminating all of the uncertainties previously referred to. The values obtained by the English group are higher for edestin, and essentially the same for egg albumin as those previously obtained. It is highly significant, however, that Chibnall found 9.8 per cent lysine in β-lactoglobulin, whereas Foster,[15] employing the isotope dilution method, found 11.4 per cent.*

TABLE 4

GLUTAMIC AND ASPARTIC ACID CONTENT OF PROTEINS DETERMINED BY VARIOUS METHODS

Protein	Glutamic Acid			Aspartic Acid		
	Isolation[14, 47]	Isotope Dilution[15]	Microbiological[52]	Isolation[14, 47]	Isotope Dilution[15]	Microbiological[53]
	per cent	per cent	per cent	per cent	per cent	per cent
Casein	22.0		21.7	6.68		7.2
β-Lactoglobulin	21.5	19.1	19.0	9.88	11.2	11.5
Egg Albumin	16.1		15.0	8.13		9.3
Gliadin	46.9		45.5	1.40		3.3
Carboxyhemoglobin (Horse)	6.75	8.5	8.9	9.3	10.3	10.8

There remain to be listed several methods which have been used in a few specific cases for the determination of individual amino acids. For example, Town[55] proposed nitranilic acid as a specific precipitant for glycine. Subsequently, however, it was shown[56] that nitranilic acid also precipitates histidine and lysine; hence, the reagent is not highly selective. Bergmann has described several specific precipitants. Potassium trioxolatochromiate[57] was employed for the determination of glycine, and accurate results were obtained in the case of gelatin[58] and silk fibroin.[59] The reagent is specific for glycine, although ammonia also forms insoluble complexes.[58] However, an empirical solubility correction must be employed, and, therefore, the reagent is probably not suited for general use, particularly when the glycine content of the hydrolysate is low. Ammonium rhodanilate has been employed for the determination of proline in gelatin,[60] and sodium dioxypyridate for the determination of alanine in silk fibroin.[59] As was subsequently shown in Bergmann's own laboratory,[58, 61] in neither case were highly accurate results obtained. The cuprous mercaptide method for cystine of Vickery and White[62] also should be mentioned. The analytical methods for the sulfur amino acids, however, will not be considered in this presentation.

---

* While this paper was in press, the second Procter Memorial Lecture of Chibnall (J. Int. Soc. Leather Trades Chem. **30**: 1. 1946) came to hand. In this lecture, Chibnall quotes newer analytical results for the basic amino acids obtained in England by Macpherson and Gale. Except in the case of arginine, these newer results are generally higher than the values referred to above. It may be pointed out that these most recent analyses were not performed exclusively by gravimetric procedure. Electrolytic, colorimetric, and gravimetric methods were employed in combination, and the decarboxylase method of Gale was employed alone or in conjunction with the electrolytic method.

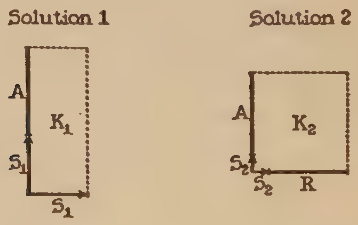
Before closing, it is pertinent to this discussion to describe briefly a method of amino acid analysis which employs specific precipitants, but which is not a quantitative isolation procedure. This method, which has been designated the solubility product method, was designed to circumvent some of the difficulties which have been cited in the procedures already discussed. Its limitations and advantages will be summarized briefly.

The solubility product method depends upon the fact that, in dilute solution, the dissociable salts of amino acids obey closely the law of the solubility product; *i.e.*, [amino acid] × [sulfonic acid] = K. An analysis by the method in its present form[63, 64] is carried out as follows (FIGURE 1):

A = mM Amino Acid in Sample

R = mM Reagent Added

S = mM. Salt Dissolved



$K_1 = S_1(A+S_1)$

$K_2 = (R+S_2)(A+S_2)$

$K_1 = K_2$: Amino acid in sample $= A = \dfrac{S_1^2 - S_2(R+S_2)}{(R+S_2) - S_1}$ Equation I

$K_1 = FK_2$: $A = \dfrac{S_1^2 - FS_2(R+S_2)}{F(R+S_2) - S_1}$ Equation II

FIGURE 1. Solubility product method.

Two equal samples of a solution are taken, each of which contains A mM of a given amino acid, say leucine. To the first sample a known amount of a sparingly soluble leucine salt is added, and the amount of the salt which has dissolved at equilibrium is determined by difference. The salt employed, which was selected on the basis of the data given in TABLE 1, was leucine 2-bromotoluene-5-sulfonate. If S_1 mM of the salt dissolved, it would contribute S_1 mM of leucine and S_1 mM of sulfonic acid to the solution, so that the solubility product in this solution would be $S_1\ (A + S_1) = K_1$. To the second sample of the unknown solution a known quantity, R mM, of the sulfonic acid is added, the quantity R being small enough so that no leucine 2-bromotoluene-5-sulfonate precipitates. The quantity of leucine salt which will dissolve in this solution is again determined as before. This quantity, S_2mM, will be less than S_1, because of the common ion effect of the sulfonic acid added. The solubility product in this second solution is $(R + S_2)(A + S_2) = K_2$. If $K_1 = K_2$, that is, if the law of the solubility product is obeyed under these experimental conditions, $S_1\ (A + S_1) = (R + S_2)(A + S_2)$. Solution of the equation for A, the only unknown, yields the desired result, namely, the amount of leucine in solution.

It will be noted that the solubility product constants, K_1 and K_2, do not appear in the final equation. The final answer, therefore, is independent of the absolute value of the solubility of the salt. This feature of the method eliminates the unknown factors concerned with the solubility of a salt in a solution of unknown composition. Amino acid salts possessing a moderate solubility in water at 0°, of the order of 1 to 5 per cent, have been found to be most advantageous to use in the method. It is thus possible to utilize reagents in this method which formerly would have been thrown into the discard. The scope of the method is thereby greatly increased.

All of the above considerations are great theoretical advantages. In practice, however, difficulties arise. For example, the solubility products K_1 and K_2 are not, in general, strictly constant. The constancy observed in the case of the leucine analysis (TABLE 5) has been shown to be the result of a compensation of several factors. In the determination of glycine with 5-nitronaphthalene-1-sulfonic acid, K_1 and K_2 were found to be unequal. By working under suitable conditions, however, the ratio K_1/K_2 can be maintained constant, and the value of this ratio can be incorporated as a factor, F, in the final equation. Mathematically, this constant ratio is as satisfactory as a constant solubility product.

The method was first checked on known solutions; and, as can be seen from TABLES 5 and 6, accurate results were obtained with both simple and

TABLE 5

l-LEUCINE ANALYSES

Reagent: Sodium Bromotoluene Sulfonate

$$\text{Leucine} = \frac{S_1^2 - S_2\,(R + S_2)}{(R + S_2) - S_1}$$

Leucine content per 3 cc.	Also present per 3 cc. in addition to N HCl		K_1	K_2	Leucine found	
mg.					mg.	Per cent recovery
12.00			9.27×10^{-4}	9.33×10^{-4}	11.74	97.8
			9.30 "	9.35 "	11.81	98.4
14.01	10% Methylcellosolve		12.10 "	12.07 "	14.13	100.8
14.90	15% Methylcellosolve		13.68 "	13.73 "	14.77	99.1
			13.72 "	13.68 "	15.04	100.9
12.13	Arginine	31.8 mg.	12.38 "	12.43 "	11.95	98.5
	Glutamic acid	26.4 "				
	Glycine	13.5 "	12.29 "	12.37 "	11.88	98.0
	Tyrosine	32.4 "				

TABLE 6

EXAMPLE OF ANALYSIS FOR GLYCINE AND *l*-LEUCINE IN A SYNTHETIC MIXTURE OF AMINO ACIDS

Percentage Composition of Mixture	
Alanine	5.8
Arginine	4.8
Aspartic acid	11.5
Cystine	5.8
Glutamic acid	11.5
Glycine	5.88
Histidine	3.5
Leucine	11.51
Lysine	3.8
Phenylalanine	11.5
Proline	11.5
Tyrosine	11.5
Ammonia	1.5

Analysis for Glycine

Found	Recovery Per cent
5.83	99.2
5.88	100.0
5.90	100.3
5.80	98.7

Analysis for *l*-Leucine

Found	Recovery Per cent
11.40	99.1
11.57	100.6
11.41	99.2

complex mixtures. Several proteins were also investigated. These proteins were chosen in order to test the method, and not with any thought that results obtained with these substances would contribute much to the problems of protein structure. It was hoped that proteins of greater interest would be analyzed later, but this was prevented by the advent of war.

Analyses of several proteins for glycine and leucine are given in TABLE 7. The results of the analyses for leucine in egg albumin and gelatin given in TABLE 7 are in good agreement with those obtained by Ryan and Brand[65] with their microbiological method (*i.e.*, 9.6 per cent leucine in egg albumin and 3.6 per cent in gelatin). Dr. Haugaard,[66] who worked in Dr. Bergmann's laboratory before the war, obtained data on β-lactoglobulin which, with his kind permission, are quoted here. Dr. Haugaard found 15.9 per cent *l*-leucine in a β-lactoglobulin hydrolysate, the figure being an average of five determinations, as follows: 15.8, 15.9, 15.2, 16.6, and 15.9 per cent. These figures were not published at the time, since the method was still in the developmental stage. The value given, however, agrees well with that of Foster (15.7 per cent), and that of Ryan and Brand (15.4 per cent).

TABLE 7

THE DETERMINATION BY THE SOLUBILITY METHOD OF GLYCINE AND *l*-LEUCINE IN SEVERAL PROTEINS

	Gm. of amino acid liberated on complete hydrolysis of 100 gm. of protein	
Protein	Glycine	*l*-Leucine
Silk Fibroin	43.3	0.96
	44.4 (43.8)*	0.75
	44.0	
Egg Albumin	3.20	9.11
	3.09	9.12
		9.00
Gelatin	25.5 (26.0)*	3.38
		3.37
Collagen	26.2 (26.5)*	3.55
(Cattle Achilles Tendon)	26.1	3.41

* Value obtained using potassium trioxalatochromiate as reagent.[58, 59]

The solubility product method, then, possesses a sound theoretical basis. Although it is not a truly micro method, analyses can be performed with a relatively small expenditure of material. Furthermore, working methods for *l*-leucine and glycine have been evolved, and apparently reliable results for these amino acids have been obtained with several protein hydrolysates.

There are, however, several factors which tend to limit the utility of the method. First among these is the lack of availability of reagents suitable for use with the method. From the list of sulfonic acids presented in TABLES 1 and 2, reagents for three or four monoamino monocarboxylic amino acids might be found. The use of the method for the estimation of the basic amino acids may be possible, but this is far from certain. The uncertainty arises from the fact that the basic amino acids form ternary salts with sulfonic acids, so that the sulfonic acid concentration occurs in the solubility product equation as a square factor. On this account, the unavoidable experimental errors will be magnified in the final result.

The fact that the solubility product is not constant means that much time and effort must be spent in the development of a method for an amino acid. The conditions have to be standardized, and the reagents highly purified, so that the ratio K_1/K_2 can be ascertained beyond question. Moreover, the quantity of hydrolysate chosen must be such that the concentration of amino acid in the unknown is nearly the same as that employed in the known. In most cases, this can be done with the aid of preliminary analyses. In some instances, however, it will not be possible, because of an unfavorable ratio between the amino acid under analysis and other amino acids which form sparingly soluble salts with the reagent being used. Although the method increases the specificity of a given reagent, it does not render the specificity absolute. Interference from other amino acids is still possible, though when it occurs it should be readily demonstrable.

The method is stereochemically specific. The determination of *l*-leucine, for example, is not affected by the presence of considerable amounts of *d*-leucine.[9] Hence, when working with hydrolysates, the occurrence of small amounts of *d*-amino acids, present in the hydrolysate as a result of racemization during hydrolysis, or for other reasons, would not be detected. If a *d*-amino acid salt is used as solid phase, the stereospecificity is reversed, and, in fact, the *d*-leucine content of gramicidin has been determined in this manner.[67] It is doubtful, however, if an analysis for one isomer can be performed in the presence of a large excess of its optical antipode.

Finally, although the method is simple in principle, accurate results are only obtained if the experimental manipulations are performed in conformity with the requirements of micro work. Slight experimental errors cause disproportionately large deviations in the final result.

In conclusion, the evidence presented in this discussion can be summarized as follows: As a result of the evaluation of the use of specific precipitants in quantitative precipitation methods, one conclusion seems inescapable. Methods of this type do not offer promise of ready attainment of the goal set at the beginning of this paper, namely, "to be able to express the empirical formulae of numerous proteins in terms of their constituent amino acids, in much the same manner as the empirical formulae of simpler compounds are expressed in terms of their constituent elements." Quantitative precipitation methods cannot be regarded as primary standard methods, since a high degree of accuracy is seldom obtained with certainty, nor can the methods be performed with the dispatch necessary. In thus assessing these methods, however, we should acknowledge our debt to them and to those who have worked with them. With these methods, information of incalculable value has already been obtained. In fact, the foundations of protein chemistry have been laid. It should occasion no surprise, however, that tools suitable for the laying of a foundation are not well suited to the fine architectural tasks entailed in completing the superstructure.

Although specific precipitants do not appear well-adapted to a primary role in the amino acid analysis of proteins, the knowledge of these reagents gained to date will be of the greatest importance. Specific precipitants remain essential tools when employed in combination with other techniques. For example, the Columbia group has employed them to great advantage in the isolations and purifications required in the course of the isotope dilution method.[15, 68, 69] The sulfonic acids, in particular, have been of great service in the routine preparation of amino acids not hitherto readily available, such as *l*-alanine, *l*-serine and *l*-phenylalanine,[11] and *l*-histidine.[70] They have also been used for the isolation of peptides from partial hydrolysates of silk fibroin.[61]

As for the solubility product method, from the point of view of general utility, the method must not only be evaluated in the light of the advantages and disadvantages inherent in the method itself, but it must also be compared with other newer methods which have developed, for the most part, since the solubility product method was first proposed. The various papers of this conference should illuminate this question.

In closing, we wish to acknowledge our indebtedness to the genius of the late Dr. Max Bergmann, in whose laboratory much of the work outlined here was performed. The product of his scientific acumen will be recognized in much that has been stated here. For any errors of omission or commission, however, the authors are, of course, solely responsible.

It is a pleasure to thank Dr. Vickery for reviewing this manuscript and offering many helpful comments.

BIBLIOGRAPHY

1. **Vickery, H. B., & C. L. A. Schmidt**
   1931. Chem. Rev. 9: 169.
2. **Mitchell, H. H., & T. S. Hamilton**
   1929. The Biochemistry of the Amino Acids. The Chemical Catalog Co. New York.

3. **Cohn, E. J., & J. T. Edsall**
   1943. Proteins, Amino Acids and Peptides. Reinhold Publishing Corp. New York.
4. **Kossel, A., & R. E. Gross**
   1924. Z. physiol. Chem. **135:** 167.
5. **Suida, W.**
   1906. Z. physiol. Chem. **50:** 174.
6. **Zimmermann, W.**
   1930. Z. physiol. Chem. **188:** 180.
7. **Zimmermann, W.**
   1930. Z. physiol. Chem. **192:** 194.
8. **Ackermann, D.**
   1934. Z. physiol. Chem. **225:** 46.
9. **Bergmann, M., & W. H. Stein**
   1939. J. Biol. Chem. **129:** 609.
10. **Doherty, D. G., W. H. Stein, & M. Bergmann**
    1940. J. Biol. Chem. **135:** 487.
11. **Stein, W. H., S. Moore, G. Stamm, C. Chou, & M. Bergmann**
    1942. J. Biol. Chem. **143:** 121.
12. **Osborne, T. B., & D. B. Jones**
    1910. Am. J. Physiol. **26:** 305.
13. **Vickery, H. B.**
    1940. J. Biol. Chem. **132:** 325.
14. **Chibnall, A. C., M. W. Rees, & E. F. Williams**
    1943. Biochem. J. **37:** 372.
15. **Foster, G. L.**
    1945. J. Biol. Chem. **159:** 431.
16. **Cannan, R. K., A. H. Palmer, & A. C. Kibrick**
    1942. J. Biol. Chem. **142:** 803.
17. **Brand, E., & B. Kassell**
    1942. J. Biol. Chem. **145:** 359.
18. **Kossel, A., & F. Kutscher**
    1900. Z. physiol. Chem. **31:** 165.
19. **Vickery, H. B., & C. S. Leavenworth**
    1926. J. Biol. Chem. **68:** 225.
20. **Vickery, H. B., & C. S. Leavenworth**
    1927. J. Biol. Chem. **72:** 403.
21. **Vickery, H. B., & C. S. Leavenworth**
    1927. J. Biol. Chem. **75:** 115.
22. **Vickery, H. B., & C. S. Leavenworth**
    1928. J. Biol. Chem. **78:** 627.
23. **Vickery, H. B., & C. S. Leavenworth**
    1928. J. Biol. Chem. **76:** 707.
24. **Osborne, T. B., C. S. Leavenworth, & C. A. Brautlecht**
    1908. Am. J. Physiol. **23:** 180.
25. **Osborne, T. B., D. D. Van Slyke, C. S. Leavenworth, & M. Vinograd**
    1915. J. Biol. Chem. **22:** 259.
26. **Kossel, A., & A. J. Patten**
    1903. Z. physiol. Chem. **38:** 39.
27. **Block, R. J.**
    1937–1938. Proc. Soc. Exp. Biol. & Med. **37:** 580.
28. **Block, R. J.**
    1934. J. Biol. Chem. **105:** 455.
    1934. Ibid. **106:** 457.

29. **Tristram, G. R.**
1939. Biochem. J. **33**: 1271.
30. **Vickery, H. B.**
1942. J. Biol. Chem. **144**: 719.
31. **Vickery, H. B., & C. S. Leavenworth**
1929. J. Biol. Chem. **83**: 523.
32. **Albanese, A. A.**
1940. J. Biol. Chem. **134**: 467.
33. **Saidel, L. J., & E. Brand**
1944 (September). Am. Chem. Soc. (New York Meeting).
34. **Dunn, M. S., M. N. Camien, S. Shankman, & L. B. Rockland**
1945. J. Biol. Chem. **159**: 653.
35. **Van Slyke, D. D., A. Hiller, & R. T. Dillon**
1942. J. Biol. Chem. **146**: 137.
36. **Drechsel, E.**
1891. Arch. Anat. u. Physiol., Physiol. Abt. **254**.
37. **Hausmann, W.**
1899. Z. physiol. Chem. **27**: 95.
1901. Ibid. **29**: 136.
38. **Van Slyke, D. D.**
1911–1912. J. Biol. Chem. **10**: 15.
39. **Van Slyke, D. D.**
1915. J. Biol. Chem. **22**: 281.
40. **Van Slyke, D. D.**
1915. J. Biol. Chem. **23**: 411.
41. **Van Slyke, D. D.**
1913–1914. J. Biol. Chem. **16**: 531.
42. **Levene, P. A., & W. Beatty**
1906. Z. physiol. Chem. **47**: 149.
43. **Zittle, C. A., & N. R. Eldred**
1944. J. Biol. Chem. **156**: 401.
44. **Dunn, M. S., M. N. Camien, S. Shankman, W. Frankl, & L. B. Rockland**
1944. J. Biol. Chem. **156**: 715.
45. **Gale, E. F.**
1945. Biochem. J. **39**: 46.
46. **Vickery, H. B., & C. S. Leavenworth**
1928. J. Biol. Chem. **79**: 377.
47. **Bailey, K., A. C. Chibnall, M. W. Rees, & E. F. Williams**
1943. Biochem. J. **37**: 360.
48. **Foreman, F. W.**
1914. Biochem. J. **8**: 463.
1914. Ibid. **8**: 481.
49. **Dakin, H. D.**
1918. J. Biol. Chem. **12**: 290.
50. **Jones, D. B., & O. Moeller**
1928. J. Biol. Chem. **79**: 429.
51. **Town, B. W.**
1928. Biochem. J. **22**: 1083.
1936. Ibid. **30**: 1837.
1941. Ibid. **35**: 417.
52. **Hac, L. R., E. E. Snell, & R. J. Williams**
1945. J. Biol. Chem. **159**: 273.
53. **Hac, L. R., & E. E. Snell**
1945. J. Biol. Chem. **159**: 291.

54. **Graff, S., D. Rittenberg, & G. L. Foster**
1940. J. Biol. Chem. **133**: 745.
55. **Town, B. W.**
1936. Biochem. J. 30: 1833.
56. **Stein, W. H., & E. G. Miller, Jr.**
1938. J. Biol. Chem. **125**: 599.
57. **Bergmann, M., & S. W. Fox**
1935. J. Biol. Chem. **109**: 317.
58. **Bergmann, M., & W. H. Stein**
1939. J. Biol. Chem. **128**: 217.
59. **Bergmann, M., & C. Niemann**
1938. J. Biol. Chem. **122**: 577.
60. **Bergmann, M.**
1935. J. Biol. Chem. **110**: 471.
61. **Stein, W. H., S. Moore, & M. Bergmann**
1944. J. Biol. Chem. **154**: 191.
62. **Vickery, H. B., & A. White**
1933. J. Biol. Chem. **99**: 701.
63. **Moore, S., & W. H. Stein**
1943. J. Biol. Chem. **150**: 113.
64. **Moore, S., W. H. Stein, & M. Bergmann**
1942. Chem. Rev. **30**: 423.
65. **Ryan, J. F., & E. Brand**
1944. J. Biol. Chem. **154**: 161.
66. **Haugaard, G.**
Unpublished data.
67. **Hotchkiss, R. D.**
1941. J. Biol. Chem. **141**: 171.
68. **Rittenberg, D., & G. L. Foster**
1940. J. Biol. Chem. **133**: 737.
69. **Shemin, D.**
1945. J. Biol. Chem. **159**: 439.
70. **Vickery, H. B.**
1942. J. Biol. Chem. **143**: 77.

THE ISOTOPE DILUTION METHOD OF AMINO ACID ANALYSIS

By David Shemin and G. L. Foster

Department of Biochemistry, College of Physicians and Surgeons, Columbia University, New York, N. Y.

There has been developed, in the last few years, a method derived from the isotope technique, which allows the estimation of amino acids in protein hydrolysates with an error which can be estimated to be within 1 to 2 per cent.[1, 2, 3, 4, 5, 6] The validity of the isotope dilution method is based on the fact that the usual laboratory procedures for isolating amino acids do not separate isotope-containing molecules from their normal analogues.

Since the isotopic compound is inseparable from its normal analogue, if one adds, for example, glutamic acid labeled with N^{15} to a protein hydrolysate and if a sample of glutamic acid is isolated from the mixture and its N^{15} excess is determined, it is possible to calculate the amount of glutamic acid (B) originally present in the protein hydrolysate from EQUATION 1,

$$B = \left[\frac{C_o}{C} - 1\right] A \qquad (1)$$

in which A is the amount of labeled glutamic acid added, and C_o and C are the values for the N^{15} excess in the added and the isolated compounds, respectively. As can be seen from the equation, the amount of the substance isolated does not enter into the calculations. All that it is necessary to know are the two isotope concentrations and the amount of glutamic acid added. For this reason, large losses are permissible during the purification procedure. This, of course, considerably simplifies the isolation of a pure specimen of the amino acid. The procedure can be used to determine any substance which can be labelled and isolated.

The errors of the isotope dilution method reside in: (A) the purity of the compound added; (B) the purity of the compound isolated; and (C) the accuracy of the isotope analyses and the value of the ratio C_o/C. Since the magnitude of these errors can be calculated and controlled, it becomes possible to estimate the accuracy of a determination.

A. Since the added amino acid can be synthesized with a very high degree of purity, no important error need arise from this source.

B. The effects of impurities in the isolated compound may be divided into three classes: (1) Contaminants which contain no nitrogen (salts, etc.) introduce no error. (2) The presence of 1 per cent of another non-isotopic amino acid introduces an error of about 1 per cent in C, which would cause an error in B of 1.25 per cent, if the C_o/C ratio is 5. However, the larger C_o/C, the smaller the error in B becomes, approaching 1 per cent as a limit. (3) Contamination of the *l* isomer with the isotopic *d* isomer will introduce a large error. In most of the determinations, the isotopic *dl* mixture was added and the diluted *l* isomer was isolated. If the isolated *l* isomer is contaminated with 1 per cent of the *d* isomer, the value of B caused by this error in C will be about 9 per cent too low, even if one works with the favorable C_o/C ratio of 10. However, in practice, the *l*-amino acid can be isolated free of the *d* component by taking advantage of the solubilities of the *l*- and *dl*-amino acids or their derivatives. Such an error, of course, cannot occur if the isotopic addendum consists of the pure *l*-amino acid.

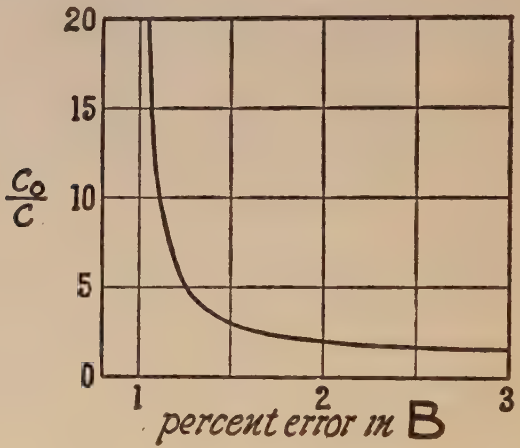


FIGURE 1. Error in result caused by a 1 per cent error in either C_0 or C at different ratios of C_0/C.

C. Above 0.5 atom per cent excess N^{15}, the mass spectrometer error of a single determination is about 1 per cent. Conditions were usually chosen so that the value for C was at least 0.5 atom per cent excess N^{15}. In FIGURE 1 is plotted the theoretical error introduced in the analysis by a 1 per cent error[1] in either C_o or C. It will be seen that little advantage is gained by increasing C_o/C above 10. In practice, this favorable ratio can usually be obtained, but in working with limited amounts of protein an unfavorable C_o/C ratio may occur with those amino acids which constitute only a small part of the protein, for it may become necessary to add comparatively large amounts of the *l*-amino acid in order to isolate enough of the pure *l*-amino acid for the N^{15} determinations at successive stages of recrystallization.

From these considerations, it can be seen that the error of the isotope dilution method is about 1 to 2 per cent.

The analytical result obtained by this method refers only to the *l*-amino acid content in the hydrolysate. Just how closely this coincides with the amount of the compound in the original protein may be somewhat uncertain, since we do not know for all of the amino acids the extent of decomposition or racemization which may occur during hydrolysis. The fact that most of the amino acids are stable, and are only very slowly racemized or decomposed when boiled with acids, does not preclude the possibility that the compounds may be more labile while in peptide linkage.

The extent of racemization can be determined by the isotope dilution method, as will be seen from the following consideration.[4] When a *dl*-amino acid containing excess N^{15} is added to the hydrolysate, the added isomers will be diluted by their isomeric counterparts in the solution. If the amino acid is subsequently isolated and fractionated into two specimens having different optical rotations, *i.e.*, containing different proportions of the two optical antipodes, the extent of the dilution of the added isomers may be computed from the specific rotation and the N^{15} concentration of the specimens isolated. The simplest case is that in which one fraction consists of the pure *l* component and the other of the optically compensated *dl* mixture. If A grams of racemic compound are added containing C_o atom per cent N^{15} excess, and the *l* and *d* isomers isolated contain C_l and C_d atom per cent N^{15} excess, respectively, then the amounts of these isomers originally present in the hydrolysate are given by EQUATIONS 2a and 2b, in which B_d and B_l are the amounts of *d*- and *l*-amino acid, respectively, in the hydrolysate.

$$B_l = \left[\frac{C_o}{C_l} - 1\right]\frac{A}{2} \tag{2a}$$

$$B_d = \left[\frac{C_o}{C_d} - 1\right]\frac{A}{2} \tag{2b}$$

On isolation of the *l* and *dl* components and determination of the N^{15} concentrations of these components, it is possible to calculate the concentration of N^{15} in the *d* isomer from EQUATION 3, in which

$$C_{dl} = \tfrac{1}{2}\, C_d + \tfrac{1}{2}\, C_l \tag{3}$$

C_{dl} is the concentration of N^{15} in the racemic mixture. Solving for C_d and substituting in EQUATION 2b, EQUATION 4 is obtained:

$$B_d = \left[\frac{C_o}{2\,C_{dl} - C_l} - 1\right]\frac{A}{2} \tag{4}$$

In those cases in which the pure *dl* mixture is not obtained, but, instead, a mixture which contains either an excess of the *l* or of the *d* isomer, it is also possible to determine the dilution of each component. From the rotation, the percentage of the *d* and *l* component can be determined, and the concentration of N^{15} in the *d* isomer can be calculated from EQUATION 5, in which P represents the fraction of the *l* enantimorph in the sample, and C_r, the concentration of N^{15} in the incompletely racemized sample.

$$C_r = PC_l + (1 - P)C_d \tag{5}$$

Solving for C_d and substituting in EQUATION 2b, EQUATION 6 is obtained:

$$B_d = \left[\frac{C_o\,(1 - P)}{C_r - PC_l} - 1\right]\frac{A}{2} \tag{6}$$

In the case in which there is no *d* component in the protein hydrolysate, the concentration of the isotope in the *dl*-amino acid isolated would be half of that in the *l* component isolated, plus half of that in the *dl* component added.

Therefore, by the use of the isotope dilution method, both the *l* and the *d* component can be estimated separately. The addition of these two values would yield a value closer to the actual amount of the amino acid in the protein.

The extent of the decomposition of a free amino acid under conditions of hydrolysis can be estimated by boiling either a protein hydrolysate or the amino acid in 6 N HCl and determining at intervals the amount of the amino acid remaining. However, if the rate of decomposition of an amino acid differs, while in peptide linkage, from that in the free state, the value obtained by any method which uses hydrolysates will differ from the actual content of the amino acid in the protein. Whether the rates of decomposition under these different conditions vary, is not as yet known.

The essential operations in the isotope dilution method, as used in this work, are: (a) hydrolysis of a weighed sample of protein for 15 to 20 hours with 20 times its weight of boiling 6 N HCl; (b) addition to the hydrolysate of a weighed amount of pure *dl*-amino acids whose N^{15} excess is known; (c) isolation from the hydrolysate of a specimen of the natural *l* isomer of the amino acid in a high state of purity; and (d) determination of the N^{15} excess in the isolated compound. In practice, several different amino acids may be determined in the same hydrolysate.

The addition of *dl* isotopic compounds to the hydrolysate necessitates the isolation of a specimen of the pure *l* isomer from the solution which contains both *l* and *d* components. In some cases, the racemic compound is less soluble than the desired *l* component and cannot be removed by recrystallization. For each different amino acid sought, some preliminary work must, therefore, be done to find a way to isolate a specimen of the *l* isomer. The procedure used to isolate pure *l*-amino acids is described briefly below.[3, 4, 5, 6] The compounds isolated from the hydrolysates were purified by recrystallization, under conditions known to lead to the pure *l* isomers. Their purity was established by the nitrogen content; the constancy of isotope concentration at successive stages of recrystallization; the specific rotation, and melting point in the cases of p-toluenesulfonylglycine, benzoylglycine, dibenzoyllysine, and the benzenesulfonyl derivatives of leucine and phenylalanine.

ANALYTICAL PROCEDURE

After the isotopic amino acids have been added to the hydrolysate, the hydrolysate is evaporated to dryness *in vacuo* and the residue dissolved in water and treated with a suspension of cuprous oxide to remove cystine and humin, as recommended by Bailey, Chibnall, Rees, and Williams.[7]

Tyrosine

The filtrate from the above procedure, after being freed of copper with hydrogen sulfide, has pH 3.5; it is concentrated for the crystallization of tyrosine. The crude tyrosine is a mixture of *l* and *dl* isomers which does not separate favorably on recrystallization. It is converted to the copper salt by being boiled with copper carbonate, and the filtrate is stored in an ice bath for crystallization of the *l* salt. The solubilities in water at 0° of *dl*- and *l*-tyrosine copper salts were found to be respectively 46 and 21 mg., calculated as tyrosine per 100 cc. of solution.[5] From the weight and the rotation of the crude tyrosine and the solubility data, one can calculate the volume required to keep all of the *dl*-tyrosine copper salt in solution while a specimen of the *l* salt crystallizes. The *l*-tyrosine copper salt is suspended in a small volume of water and treated with an equivalent amount of HCl. The tyrosine so obtained has the correct rotation[8] and is recrystallized two or three times, a sample being saved at each stage for N^{15} analysis.

Leucine

The filtrate from the crude tyrosine is adjusted to about pH 6.0, concentrated to a small volume, and left several days in the cold. From the "crude leucine" fraction, *l* leucine is isolated as its salt with 2-bromotoluene-5 sulfonic acid,[9] the solubilities of the *l* and *dl* salts in 1 N HCl at 0° being about 1.2 and 4.1 gm., respectively, per 100 cc. of solution.[5] The leucine is purified as the benzenesulfonyl derivative which has very favorable properties; *e.g.*, m.p. 119°, $[\alpha]_D^{22} = -53°$ (in the presence of 6 moles of NaOH and $C = 2$), solubility in water at 0°, *l* form 47 mg., *dl* form 118 mg. per 100 cc. The leucine salt of bromotoluene sulfonic acid is benzenesulfonated directly. Leucine appears to be stable in boiling 6 N HCl, since the values obtained for leucine from hemoglobin hydrolysates boiled for 20 and 94 hours were the same.[5]

Phenylalanine

If the protein is not too poor in phenylalanine, this amino acid will be present in the "crude leucine" fraction and will pass into the filtrate from the leucine salt, from which it may be isolated as its salt with 2,5-dibromobenzenesulfonic acid.[9] Final purification of the *l*-phenylalanine is effected by recrystallization of its benzenesulfonyl derivative from hot water; m.p. 133°, $[\alpha]_D^{22} = -35.5°$ (in the presence of 6 moles of NaOH and $C = 2$); solubilities in water at 0°, *l* form 37 mg., and *dl* form 85 mg. of the sulfonyl derivative per 100 cc. of solution.[5]

The possibility that phenylalanine is racemized during hydrolysis has been raised by Martin and Synge.[10] In order to test this point, 0.800 gm. of *dl*-phenylalanine having 1.71 per cent N^{15} excess was added to an acid hydrolysate of 31.4 gm. of horse hemoglobin. The *l*-phenylalanine isolated had 0.264 per cent N^{15} excess, while the *dl*-phenylalanine isolated had 0.904 per cent N^{15} excess. From these data, it was calculated that the original hydrolysate contained 2.16 gm. of *l*-phenylalanine and 0.04 gm. of *d*-phenylalanine derived from the protein. Thus, only 1.8 per cent of the total phenylalanine from the protein was the *d* isomer.[5] This is about the magnitude of the error to be expected in the experiment, and is not significant.

That there is destruction of phenylalanine in boiling 6 N HCl is indicated by the finding of ammonia and the odor of phenylacetaldehyde in the reaction mixture. An experiment designed to measure by isotope dilution the extent of this decomposition of phenylalanine in boiling 6 N HCl solution indicated a destruction of about 5.5 per cent of the compound in 40 hours. Hence, it seems likely that the reported figure for phenylalanine in horse hemoglobin is 2 or 3 per cent too low.[5]

Arginine

From the filtrate from the "crude leucine" fraction, arginine is isolated as the monoflavianate, in the usual way. The monoflavianate is recrystallized from 0.1 N HCl, with a volume sufficient to keep in solution the amount of *dl*-arginine that was added to the hydrolysate. The solubilities of *l*- and *dl*-arginine monoflavianates in 0.1 N HCl at 25° were found to be 29 and 56 mg., respectively, per 100 cc. The *l*-arginine monoflavianate is converted to the di-3,4-dichlorobenzenesulfonate[9] by dissolving the monoflavianate in hot 1 N HCl and adding a large excess of dichlorobenzenesulfonic acid. This salt is recrystallized several times. The anhydrous salt has $[\alpha]_D^{23} = +8.0°$ ($C = 6$ in 95 per cent alcohol); the solubilities of the *l*- and *dl*-arginine salts in 1 N HCl at 0° were, respectively, 117 and 112 mg. per 100 gm. of solution.[5]

Lysine

The filtrate from arginine flavianate is made strongly acid with HCl and shaken with butanol to remove the excess flavianic acid. The excess of HCl is removed by distillation, and histidine is precipitated in the usual way with $HgCl_2$ at pH 6.8.[11] After the removal of mercury with H_2S, lysine is precipitated with phosphotungstic acid. Lysine is recovered from the phosphotungstate as the monohydrochloride, which is converted to the dibenzoyl derivative and recrystallized from acetone until it has $[\alpha]_D^{22} = -8.6°$ ($C = 5$ in methyl alcohol), and a melting point of 149.5–150°.

Lysine was also isolated as the ε-monobenzoyl derivative, according to the method of Kurtz,[12] from a protein hydrolysate from which cystine, tyrosine, and the leucine-phenylalanine fraction were first removed. The purity of the ε-benzoyl-*l*-lysine was established by nitrogen analysis and by constancy of isotope concentration and specific rotation on successive recrystallization; N 11.1 per cent (calculated 11.2 per cent); $[\alpha]_D^{22} = +18.8°$ (2.3 per cent in N HCl).[6] The specific rotation in N HCl of the isolated ε-benzoyl-*l*-lysine was constant during five recrystallizations and was the same as that of a sample prepared from pure *l*-lysine. The value observed for the reference sample in 50 per cent acetic acid was +11.4°, in contrast to +27.2° reported by Goldschmidt and Kinsky.[13] The reason for this discrepancy is not clear.

Glutamic Acid

The filtrate from lysine phosphotungstate is freed from phosphotungstic acid by extraction with a butanol-ether mixture. The dicarboxylic acid fraction is obtained in the usual way with barium hydroxide and alcohol. From this fraction, glutamic acid is isolated and purified as the hydrochloride, the solubilities of the *l* and *dl* hydrochlorides in 5.4 N HCl at 0° being 1.24 gm. and 2.48 gm. per 100 gm. of solution, respectively.[4] It has been found by the method outlined above that the amount of *d*-glutamic acid existing in tissue protein hydrolysates does not exceed 1 per cent.[4]

The necessity of removing cystine prior to the precipitation of the barium dicarboxylates was demonstrated by the following observation.[6] In the purification of glutamic acid hydrochloride from a hydrolysate of human serum albumin which had not been treated with cuprous oxide, the isotope content, at successive stages of recrystallization, was constant (0.736, 0.719, 0.735 atom per cent N^{15} excess), and the nitrogen content was 7.7 per cent, from which it might have been concluded that the substance was pure *l*-glutamic acid hydrochloride. However, the specific rotation, even after five recrystallizations, was —4.3°, instead of +31°. The contaminant proved to be cystine dihydrochloride. It was found that the solubility of cystine dihydrochloride in 6.05 N HCl at 0° is considerably lower than that of glutamic acid hydrochloride: 0.51 gm. and 0.90 gm. per 100 gm. of solution of cystine dihydrochloride and glutamic acid hydrochloride, respectively.[6] This observation demonstrates the possibility of failure to detect contamination of one amino acid with another if the purity is tested merely by nitrogen determination, or by carbon and hydrogen analysis only. Contamination of glutamic acid with cystine would not be detected by a carbon analysis, unless the contamination were greater than 4 per cent, and would not be detected by nitrogen analysis, unless the cystine content were greater than 15 per cent.

Aspartic Acid

The filtrate from the glutamic acid hydrochloride is taken to dryness *in vacuo*. From the residue, aspartic acid is isolated as the copper salt. *l*-Aspartic acid is regenerated from the recrystallized copper aspartate and repeatedly recrystallized from water. The specific rotation, based on free aspartic acid, is $[\alpha]_D^{25} = +25.4°$ (1.5 to 3 per cent in 1 to 2.5 N HCl).

Glycine

From the filtrate from the barium salts of the dicarboxylates, freed from barium and alcohol and concentrated *in vacuo,* glycine is isolated as the trioxalatochromiate, by the method of Bergmann and Fox,[14] and purified either as the p-toluenesulfonyl derivative, m.p. 147–148°, or as the benzoyl derivative, m.p. 190°. The values reported for glycine are not as precise as those for the other amino acids, for the ratio of the isotope concentration of the added glycine (C_o) to the isotope concentration of the isolated glycine (C) is low. As the proteins analyzed contain but little glycine, it was necessary to add comparatively large amounts of glycine, in order to isolate enough of the pure amino acid for the N^{15} determinations at successive stages of recrystallization.

The possibility that glycine might be formed by a decomposition of serine in acid solution, analogous to that known to occur in an alkaline medium,[15] was investigated by the isotope dilution method. To 990 mg. of serine, 500 mg. of isotopic glycine, containing 1.095 atom per cent N^{15} excess, were added, and the mixture was refluxed for 17 hours with 100 cc. of 20 per cent HCl. From this mixture, glycine was precipitated with trioxalatochromiate and purified as the p-toluenesulfonyl derivative. The isotope concentration of the isolated glycine was found to be 1.091 atom per cent N^{15} excess. Since no significant dilution of the added glycine had occurred, it was concluded that none of the glycine in the protein hydrolysate was derived from serine.[6]

A summary of the analytical data and results obtained for the above-mentioned *l*-amino acids, in the following protein hydrolysates: cattle fibrin, horse hemoglobin, β-lactoglobulin, bovine serum albumin and human albumin, is given in TABLES 1–6.[3, 4, 5, 6]

TABLE 1

DETERMINATION OF AMINO ACIDS IN FIBRIN[3]

Experiment No.	Weight of protein hydrolyzed (D)	Amino acid added: Compound	Amino acid added: Amount of *l*-amino acid (A)	Amino acid added: N^{15} excess (C_o)	N^{15} excess in compound isolated at successive stages of recrystallization (C)	Amino acid in protein $\left(\frac{C_o}{C}-1\right)\times\frac{100A}{D}$
	gm.		*gm.*	*atom per cent*	*atom per cent*	*per cent*
17	4.002	Glycine	0.1315	2.00	0.781 0.784 0.782 0.782	5.12
21	4.995	*l*-Glutamic acid	0.0633	2.10	0.182 0.186 0.185	13.2
22	5.382	" "	0.0576	2.10	0.162 0.161 0.157	13.0
23*	5.034	" "	0.0611	2.10	0.185 0.176	12.9
25†	7.675	" "	0.2003	2.10	0.356 0.355	12.8
		l-Aspartic acid	0.2000	2.00	0.374 0.383	11.2

* In this experiment, the isotopic glutamic acid was added before the hydrolysis was begun.
† In this experiment, both glutamic and aspartic acids were added to the same hydrolysate.

TABLE 2

ANALYSES OF HORSE HEMOGLOBIN[5]

Experiment No.	Weight of protein hydrolyzed (D)	Amino acid added			N^{15} excess in compound isolated at successive stages of recrystallization (C)	Amino acid in protein $\left(\frac{C_o}{C}-1\right)\times\frac{100A}{D}$
		Compound	Amount of *l*-amino acid (A)	N^{15} excess (C_o)		
	gm.		*gm.*	*atom per cent*	*atom per cent*	*per cent*
4	21.65	Tyrosine	0.0740	6.85	0.694, 0.693, 0.710	3.0
5	25.60	"	0.0734	6.85	0.602, 0.602	3.0
7	24.58	"	0.0768	6.85	0.645, 0.654, 0.647	3.0
4	21.65	"	0.0739	6.79	0.324, 0.320, 0.327	6.8
5	25.60	"	0.1296	6.79	0.466, 0.478, 0.472	6.8
7	24.58	"	0.1522	6.79	0.552, 0.563, 0.563	6.9
1	21.60	Arginine	0.0850	8.72	0.845, 0.831	3.7
4	21.65	"	0.0710	8.72	0.722, 0.697	3.7
5	25.60	"	0.0614	8.72	0.530, 0.527, 0.529	3.7
7	24.58	Lysine	0.1372	9.09	0.559, 0.557	8.5
8	22.13	"	0.1542	9.09	0.674, 0.667, 0.671	8.7
11	20.39	"	0.1430	9.09	0.690, 0.691, 0.691	8.5
5	25.60	Glutamic acid	0.2030	4.52	0.383, 0.384, 0.381	8.6
7	24.58	" "	0.2211	4.52	0.437, 0.440, 0.435	8.4
8	22.13	" "	0.2446	4.52	0.528, 0.528, 0.528	8.4
8	22.13	Aspartic acid	0.1642	6.75	0.452, 0.449, 0.450	10.4
11	20.39	" "	0.1420	6.75	0.421, 0.426	10.4
8	22.13	Glycine	1.1869	1.15	0.570, 0.574, 0.565	5.5
11	20.39	"	0.3498	3.38	0.791, 0.789, 0.787	5.6
2	4.85	Leucine	0.0767	6.73	0.646, 0.642, 0.626	15.1
3	4.69	"	0.0621	6.73	0.547, 0.552, 0.537	15.0
6	2.44	"	0.0300	6.73	0.507, 0.508	15.1
10*	1.65	"	0.0210	6.73	0.525	15.0

* Hydrolysis continued for 94 hours.

TABLE 3

ANALYSES OF β-LACTOGLOBULIN[5]

Lacto-globulin specimen	Weight of protein hydro-lyzed (D)	Amino acid added			N^{15} excess in compound isolated at successive stages of recrystallization (C)	Amino acid in protein $\left(\frac{C_o}{C}-1\right)\times\frac{100A}{D}$
		Compound	Amount of *l*-amino acid (A)	N^{15} excess (C_o)		
	gm.		*gm.*	*atom per cent*	*atom per cent*	*per cent*
I	7.05	Glutamic acid	0.1484	4.52	0.448, 0.448	19.1
II	6.37	“ “	0.1413	4.52	0.470, 0.476	19.0
I	7.05	Aspartic acid	0.0681	6.75	0.538, 0.532, 0.532	11.3
II	6.37	“ “	0.0624	11.82	0.947, 0.961	11.2
I	7.05	Lysine	0.0660	9.09	0.689, 0.683, 0.687	11.4
II	6.37	“	0.0641	9.09	0.736, 0.736	11.4
I	7.05	Leucine	0.0688	6.73	0.392, 0.393, 0.403	15.6
II	6.37	“	0.0698	6.73	0.436, 0.441	15.7
I	7.05	Glycine	0.3147	1.99	1.474, 1.496	1.5
II	6.37	“	0.2365	2.42	1.73, 1.73	1.5

TABLE 4

ANALYSIS OF CRYSTALLINE BOVINE SERUM ALBUMIN[6]

Experiment No.	Weight of protein hydrolyzed (D)	Amino acid added			N^{15} excess in compound isolated at successive stages of recrystallization (C)	Amino acid in protein $\left(\frac{C_o}{C}-1\right) \times \frac{100A}{D}$
		Compound	Amount of *l*-amino acid (A)	N^{15} excess (C_o)		
	gm.		*mg.*	*atom per cent*	*atom per cent*	*per cent*
1	4.925	Glutamic acid	39.9	18.69	0.840, 0.857, 0.866	16.9
4	9.10	" "	94.7	18.69	1.091, 1.071, 1.079	17.0
1	4.925	Aspartic acid	36.75	11.82	0.790, 0.800	10.4
4	9.10	" "	90.60	11.82	1.057, 1.051, 1.056	10.2
9	7.12	Tyrosine	63.37	6.85	0.954, 0.949, 0.945	5.5
1	4.925	Glycine	307.1	3.13	2.38, 2.43, 2.37	2.0
4	9.10	"	300.7	3.01	1.92, 1.87, 1.87	2.0
9	7.12	Lysine	49.89	9.09	0.483, 0.488, 0.485	12.4

TABLE 5

ANALYSIS OF CRYSTALLINE HUMAN SERUM ALBUMIN[6]

Experiment No.	Amount of protein hydrolyzed (D)	Amino acid added			N^{15} excess in compound isolated at successive stages of recrystallization (C)	Amino acid in protein $\left(\frac{C_0}{C}-1\right)\times\frac{100A}{D}$
		Compound	Amount of *l*-amino acid (A)	N^{15} excess (C_0)		
	gm.		*mg.*	*atom per cent*	*atom per cent*	*per cent*
2	10.49	Glutamic acid	90.16	18,69	0.891, 0.884, 0.891	17.2
3	7.80	" "	89.37	18.69	1.183, 1.185, 1.187, 1,195	16.9
5	7.98	" "	86.80	18.69	1.115, 1.132, 1.136, 1.133, 1.139	16.9
7	5.86	" "	73.72	18.69	1.268, 1.281, 1.294	17.1
2	10.49	Aspartic acid	84.20	11.82	0.873, 0.873, 0.876	10.0
3	7.80	" "	91.05	11.82	1.274, 1.275, 1.267, 1.261	9.7
5	7.98	" "	74.60	11.82	1.047, 1.064, 1.055, 1.053	9.6
5	7.98	Tyrosine	77.65	6.85	1.160, 1.158	4.8
7	5.86	"	379.8	1.025	0.596, 0.595, 0.590, 0.600	4.7
5	7.98	Glycine	266.3	3.03	2.06, 2.05, 2.05, 2.05	1.6
7	5.86	"	352.7	2.54	1.99, 2.06, 2.05, 2.06	1.5

TABLE 6

AMINO ACID CONTENT OF HYDROLYSATES OF HORSE HEMOGLOBIN, β-LACTOGLOBULIN, BOVINE AND HUMAN SERUM ALBUMIN

					Number of Residues per Mole			
Amino Acid	Hemoglobin	β-Lactoglobulin	Bovine Serum Albumin	Human Serum Albumin	Hemoglobin M.W. = 66,700	β-Lactoglobulin M.W. = 42,000	Bovine Serum Albumin M.W. = 70,000	Human Serum Albumin
	per cent	*per cent*	*per cent*	*per cent*				
Glutamic acid	8.5	19.1	17.0	17.0	39	55	81	81
Aspartic acid	10.3	11.3	10.3	9.8	52	36	54	52
Leucine	15.1	15.7			77	50		
Glycine	5.6	1.5	2.0	1.6	50	8	18	15
Phenylalanine	6.8				27			
Tyrosine	3.0		5.5	4.7	11		21	18
Arginine	3.7				14			
Lysine	8.6	11.4	12.4		39	33	60	

BIBLIOGRAPHY

1. **Schoenheimer, R., S. Ratner, & D. Rittenberg**
   1939. J. Biol. Chem. **130**: 703.
2. **Ussing, H. H.**
   1939. Nature **144**: 977.
3. **Rittenberg, D., & G. L. Foster**
   1940. J. Biol. Chem. **133**: 737.
4. **Graff, S., D. Rittenberg, & G. L. Foster**
   1940. J. Biol. Chem. **133**: 745.
5. **Foster, G. L.**
   1945. J. Biol. Chem. **159**: 431.
6. **Shemin, D.**
   1945. J. Biol. Chem. **159**: 439.
7. **Bailey, K., A. C. Chibnall, M. W. Rees, & E. F. Williams**
   1943. Biochem. J. **37**: 360.
8. **Stein, W. H., S. Moore, & M. Bergmann**
   1942. J. Am. Chem. Soc. **64**: 724.
9. **Doherty, D. G., W. H. Stein, & M. Bergmann**
   1940. J. Biol. Chem. **135**: 487.
10. **Martin, A. J. P., & R. L. M. Synge**
    1941. Biochem. J. **35**: 91.
    Gordon, A. H., A. J. P. Martin, & R. L. M. Synge
    1943. Biochem. J. **37**: 39.
11. **Foster, G. L., & D. Shemin**
    1938. Organic Syntheses **18**: 43. New York.
12. **Kurtz, A. C.**
    1941. J. Biol. Chem. **140**: 705.
13. **Goldschmidt, S., & A. Kinsky**
    1929. Z. physiol. Chem. **183**: 244.
14. **Bergmann, M., & S. W. Fox**
    1935. J. Biol. Chem. **109**: 317.
15. **Daft, F. S., & R. D. Coghill**
    1943. J. Biol. Chem. **90**: 341.

CHROMATOGRAPHIC AND ION EXCHANGE METHODS OF AMINO ACID ANALYSIS

By R. Keith Cannan

New York University College of Medicine, New York, N. Y.

The procedures which will be discussed under this title are those which depend upon the differential distribution of amino acids between two phases, using some form of countercurrent operation. Classical examples of countercurrent separations of amino acids are the fractional distillation of the esters, the continuous extraction of the neutral amino acids from aqueous solution by butyl alcohol, and the separation of the acidic and basic amino acids by electrodialysis. In none of these procedures, however, have the potentialities of countercurrent operation been efficiently exploited and none has been refined to the point of achieving the separation of individual amino acids from one another. In the past few years, interest has been aroused in the application to mixtures of amino acids of the methods of chromatographic analysis which have proved so useful in other fields. Attention has been given, not only to the use of solid adsorbents, but also to the differential distribution of amino acids and their derivatives in liquid-liquid systems. It is to this recent work that this review will be confined.

Adsorption analyses, as is well known, are usually carried out by perfusing a solution of the solutes which are to be analyzed through a column of the finely-divided adsorbent. In a successful analysis, the solutes are adsorbed as contiguous bands on the column. If the column is then perfused with the solvent, the bands advance at different rates and become discretely separated. The individual solutes may then be recovered by sectioning the column and extracting the several sections. More simply, the intact column may be perfused with the solvent or a series of solvents, until each solute appears in turn as a single component in the effluent from the column.

The application of the countercurrent principle to extractions in liquid-liquid systems would appear to require a multichamber extraction assembly, through which one or both of the liquids could be continuously circulated. Technical difficulties in designing a compact apparatus, in which equilibrium conditions can be maintained during continuous operation, have restricted the use of this type of analysis. Attention may be drawn, however, to an interesting assembly which has been designed by Craig[1] for discontinuous operation. This apparatus promises to be most useful in experiments in which it is desirable to make intimate examinations of the distribution of solutes in any or all of the chambers, at any stage of the analysis. Meanwhile, the problem of continuous extraction has been reduced to a simple process of column filtration, by Gordon, Martin, and Synge.[2] They accomplish this by the elegant device of immobilizing one liquid in the interstices of a column of an inert gel or fibrous reticulum, through which the second solvent is perfused. The solutes appear as separate bands in the column, and these bands behave in essentially the same manner as do those formed in an adsorption column. The authors have used the method extensively in the analysis of mixtures of amino acids. It will, certainly, have many other applications.

Methods of adsorption and partition analysis, carried out in columns, have much to recommend them to the attention of those concerned with the analysis of proteins. In the first place, the methods are particularly suited to the separation of substances which are alike in their reactive groups, but differ in their non-polar characteristics. They should be, and, indeed, have already proved themselves to be, particularly valuable in the analysis of mixtures of neutral amino acids. Secondly, the methods not only are adaptable to small amounts of material, but also permit several components to be analyzed in a single sample. Moreover, the fractions are not spoiled for subsequent analysis by the introduction of any foreign reagent. For this reason, the fractions which are obtained should be very suitable for ultimate analysis by the method of isotope dilution. Finally, the process of serial elution of components is one which lends itself to methods of graphic analysis of the composition of the sample, based upon continuous records of the concentration of solutes in the eluate. Continuously recording refractometers, interferometers, and potentiometers have been used for this purpose.

THEORETICAL CONSIDERATIONS

Although chromatography has been practised by analysts for many years, the first attempt to develop a quantitative theory was not published until 1940. The treatment leads to a comprehensive description of the behavior of a single solute, but the case of a mixture of solutes has defied a general solution. Even so, the results do account for a number of the vagaries of chromatograms, and merit brief review. In order to include both adsorption and partition chromatography in the discussion, we will refer to the phase which constitutes the column as the *stationary phase,* and to the perfusing fluid as the *mobile phase*. The transfer of solute from the mobile to the stationary phase will be described as an absorption.

The following symbols will be used:

c — equilibrium concentration of solute in the mobile phase,
c_0 — initial concentration of solute in the mobile phase,
q — amount of solute absorbed per unit length of column,
M — mass (volume) of stationary phase per unit length of column,
$q = Mf(c)$ — absorption isotherm,
$f'(c)$ — first derivative of $f(c)$,
w — pore volume per unit length of column,
x — distance of a point from the origin of the column,
x_r — length of column,
v_0 — volume of solution used to form the band,
v — volume of solvent used to develop the band,
$v_r = (v + v_0 - x_r w)$ — retardation volume for the case $x = x_r$,
$\overline{v}_r$ — specific retardation volume $= v_r/Mx_r$.

The theory is based on the assumption of essential equilibrium in all elements of the column. The effects of diffusion on the distribution of solute in the band are ignored. The differential equation for these conditions was first derived by Wilson,[3] who considered some of its implications. It was developed in greater detail, independently by De Vault[4] and by Weiss,[5] both of whom have emphasized the importance of the nature of the function, $f(c)$, in determining the behavior of a chromatogram. The equation is based on the conservation condition for the absorption process in an element of the column of length dx (containing solution in equilibrium with it), when an element of solution of volume dv and concentration c, passing to $(c - dc)$, is admitted. The assumption of equilibrium leads to the equation:

$$\left(\frac{dv}{dx}\right)_c = w + Mf'(c). \qquad (1)$$

This is the differential equation of a straight line in variables v and x. The general solution is a family of straight lines of slopes $w + Mf'(c)$ and positions determined by the initial conditions. If the column is initially empty, the relation of v to x, for values of c between c_0 and 0, reduces to the simple expression:

$$v/x = w + Mf'(c). \qquad (2)$$

EQUATION 2 defines the volume flow of mobile phase relative to the movement of an element of concentration, c, in this phase. Since the absorbed band is in equilibrium with the mobile phase, the movement of the band is also defined. In the discussion which follows, it will be convenient to refer to $v/x - w$, or, neglecting w, to v/x, as the retardation of the element of the band which is under consideration.

The function, $f(c)$, may assume a variety of forms, of which three types may be distinguished:

A. $f(c)$ is proportional to c,
B. $f(c)$ increases less rapidly than c,
C. $f(c)$ increases more rapidly than c.

A few examples of these types are given in TABLE 1. It should be noted that, in the case of a liquid-liquid partition, the typing of the isotherm depends upon which of the two solvents constitutes the stationary phase. By reversing the roles of the two liquids, a Type B system becomes a Type C, and vice versa. In TABLE 1, k, in each case, is the Mass Action constant for the secondary reaction in question. The constants, a, represent the value of the distribution coefficients, $\frac{q}{Mc}$, when the secondary reactions are negligible.

TABLE 1

$q = Mf(c)$

Type	f(c)	f′(c₀)	f′(o)	Condition
		Adsorption Isotherms		
A	ac	a	a	Linear Isotherm, *e.g.*, adsorption from very dilute solutions.
B 1	$ac/(1 + kc)$	$a/(1 + kc_o)^2$	a	Langmuir Isotherm.
		Partition Isotherms		
A	ac	a	a	Linear partition.
B 2	$ac + \sqrt{kac}$	$a + 0.5\sqrt{\frac{ka}{c_o}}$	α	Ionization of solute in the stationary phase.
B 3	$ac + 0.5\sqrt{kac}$	$a + 0.25\sqrt{\frac{ka}{c_o}}$	α	Dissociation of solute dimer in the stationary phase.
C	$ac + 2k(ac)^2$	$a(1 + 4kac_o)$	a	Dimerization of solute in the stationary phase.

With the aid of TABLE 1, EQUATION 2 may be used to determine the effect on a typical absorption band of passing solvent through it. Let the band be represented by a zone of uniform concentration in equilibrium with c_o. At the two boundaries, there will be a concentration gradient, $c_o \rightarrow o$. Consider the effect on these boundaries of the passage of solvent through the band. The band will dissolve at the rear face and reform at the forward edge. That is to say, the boundaries will advance in the direction of the flow of solvent, but at a retarded rate. In the case of a linear isotherm, the boundaries will retain their original gradient, because, $Mf'(c)$ being independent of c, all regions of the boundaries will be equally retarded. In the case of an isotherm of Type B, on the other hand, $Mf'(c)$ diminishes as c increases, with the result that the forward boundary will tend to sharpen,* and the trailing edge to become more diffuse. At the same time, the region of uniform concentration will shrink, as the zone of varying concentration in the tail progressively extends forward toward the leading boundary. The band is said to be fully developed when no region of uniform concentration, in equilibrium with c_o, remains.

It is interesting to note that, for two of the isotherms given in TABLE 1 (B 2 and B 3), $Mf'(o)$ is infinite. This implies that the origin of the rear boundary should not move from its original position. In practice, it probably means that solutes of this type will be difficult to separate from others or to elute completely.

The situation, in the case of Type C systems, is the reverse of that for Type B. The band will develop diffuseness at the front and have a sharp trailing boundary.

The effects of development on bands of Types B and C are illustrated in FIGURE 1. The initial boundaries are assumed to be sharp. The conditions of partial and full development are illustrated. The broken lines (shown only for Type B) represent the discontinuities given by the theory. The progress of events in the course of the formation, development, and elution of the chromatogram of a single solute will now be reviewed.

Formation

A volume, v_o, of solution of concentration, c_o, is admitted to a column containing no mobile phase. Successive elements of the stationary phase absorb solute until they reach equilibrium at $c = c_o$. A band of uniform density develops. In the absence of diffusion, the boundary of a system of

---

* Theoretically, a discontinuity develops, as shown in the broken lines of FIGURE 1. Physically, this means that the band will tend to maintain a sharp front against the forces of diffusion, etc.

Type A or B will be sharp. Let its position be x_o. The conservation of matter requires that

$$(v_o - x_o w)c_o = x_o q = x_o M f(c_o). \tag{3}$$

That is to say,

$$\bar{v}_r = \frac{v_o - x_o w}{x_o M} = \frac{f(c_o)}{c_o} = \underset{\text{(linear isotherm)}}{a} = \underset{\text{(adsorption isotherm)}}{\frac{a}{1 + kc_o}}. \tag{4}$$

The quantity, $\bar{v}_r$, has been called the specific retardation volume of the solute, by Tiselius.[6] It represents the volume of solution which would be

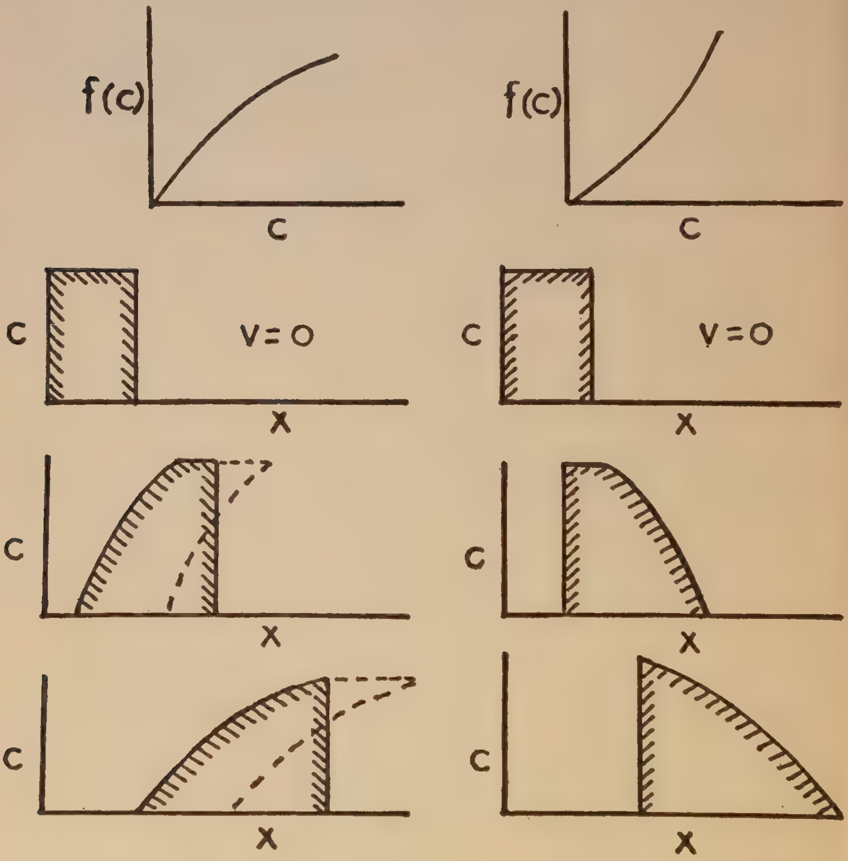


FIGURE 1.

cleared of solute by a column of unit length, containing unit mass (volume) of stationary phase. It is conveniently measured by recording the volume, v_r, of solute free solvent which appears as effluent up to the moment that the band reaches the end of the column. Then, $v_r/Mx_r = \bar{v}_r$. Attention is drawn to the fact that, whereas $\bar{v}_r$ is independent of c_o in the case of a linear isotherm, its value may be expected to diminish as c_o increases, in the majority of adsorption analyses and in liquid-liquid partitions of Type B.

The measurement of the retardation of a solute of Type C is complicated by the diffuseness of the leading boundary. The volume of effluent leaving the column, up to the moment of the appearance of the first trace of solute, should, theoretically, be equal to $Mf'(o) = aM$ and should be independent of c_o. On the other hand, the volume collected, up to the time that the concentration in the effluent becomes equal to c_o, should be equal to $Mf'(c_o) = Ma(1 + 4\,kac_o)$, for the case quoted in TABLE 1. It might prove difficult, in practice, to make precise measurement of either of these volumes.

Development

When the volume, v_o, of solution is succeeded by a volume, v, of pure solvent, the absorbed band is progressively desorbed at the rear and reformed at the front. The band detaches itself from the origin of the column and advances. The changes which take place in the boundaries of the different types of band have already been considered. The forward boundaries of systems of Types A and B are sharp. Their retardations are given by EQUATION 4, where

$$\bar{v}_r = \frac{(v + v_o) - x_r w}{x_r M} = \frac{f(c_o)}{c_o}. \qquad (5)$$

That is to say, the same retardation volume should be obtained when a limited band is formed and developed (Elution Analysis), as that which is obtained when the perfusion with solution is continued until solute appears in the effluent (Front Analysis). This is true, however, only for partially developed bands. If the band becomes fully developed before elution, the equilibrium concentration at the leading edge begins to fall below c_o — a change which will tend to raise $\bar{v}_r$ above that given by EQUATION 5.

Similar considerations apply to the retardation volume measured for the sharp rear boundary of a band of Type C. For a partially developed band of the type cited in TABLE 1,

$$\bar{v}_r = \frac{v - x_r w}{x_r M} = a(1 + 2ka^2c_o). \qquad (6)$$

Elution

The curve relating the volume flow of effluent to the concentration of solute will reflect the character of the band which is being eluted. A linear isotherm should give an eluate of uniform concentration equal to c_o. Other types will give concentration curves, dependent on the types of isotherm which apply and the degree of development of the bands at the time of elution.

A solute may be so strongly retarded in the original solvent that it becomes profitable to develop and elute the band with a different solvent. In this case, the retardation volume and the concentration of the eluate are determined by the isotherm for the new solvent. As an alternative to a change of solvent, the chromatogram may be developed by displacement—*i.e.*, with a solution of a more readily absorbed substance. This method implies a degree of competition between the two solutes and is, probably, of significance only in adsorption analysis. The simplest case, that of an exchange reaction, will be considered in a later section of this review. It should be noted that the bands remain contiguous during development by displacement; for, in effect, each band advances because it is pushed forward by the one behind it.

Multiple Solutes

Consider the case of two solutes, A and B, B being more strongly absorbed than A. During formation, the chromatogram consists of two bands. The band distal from the origin of the column will contain A alone, while the proximal band will contain both solutes in proportions determined by their adsorption coefficients in the mixture. If neither solute affects the distribution of the other, the boundary of each solute will behave as though the other solute were absent. In general, however, the distribution of each solute may be expected to be affected by the presence of the other. In a partition analysis, the result of solute interaction may be either to inhibit or to favor absorption by the stationary phase. When the latter is a solid adsorbent, on the other hand, the effect of a second solute may be expected to inhibit the adsorption of both. The retardations of

A and B, in the mixed band, will both be less than for the same concentrations of pure components. It follows that there will be a continual tendency to feed A forward to the leading band. In other words, the equilibrium concentration of A in the latter band will be greater than in the original solution. The following treatment of the condition for a Front Analysis of A is essentially that of Tiselius.

Let v_A and v_B be the retardation volumes corresponding to the first appearance of A and B, respectively, in the effluent from the column, and let c_A and c'_A be the concentrations of A in the influent and effluent, respectively. Then, if $Mf(c_A')$, in the absence of B, be equal to ac'_A, and $Mf(c_A)$, in the presence of B, be equal to $a'c_A$, we have the conservation condition that $v_B c_A = (v_B - v_A)c'_A + Mx_r a' c_A$. Also, $v_a c'_A = Mx_r a c'_A$. Combining and rearranging, we have $\frac{c'_A}{c_A} = \frac{v_B - Mx_r a'}{v_B - Mx_r a}$. If a′ is smaller than a, it is evident that c_A' will be larger than c_A.

In the process of development, the mixed band will detach itself from the origin of the column and, as A continues to feed forward into the first band, a band of pure B will develop in the rear. Four boundaries must, therefore, be considered. The first and fourth boundaries are those of single components. Thus, the leading one may be expected to be sharp, and the trailing boundary, diffuse. In the intermediate boundaries, the concentration of one component is also changing from c_o to o, but this condition is complicated by a simultaneous increase in the concentration of the other solute. Consequently, although the leading boundary of the mixed band may be expected to be sharp and its trailing boundary diffuse, this need not, necessarily, be so. De Vault[4] has made a careful analysis of the complex changes which may occur in this stage of development.

Under favorable conditions of development, the mixed band will ultimately disappear and the two solutes will then appear as separate bands, which will move independently, at rates determined by their adsorption isotherms and the concentrations (differing from the original concentrations) which are maintained at the respective band fronts. Assuming the solutes are of type B, they should maintain or develop boundaries which are sharp at the front and diffuse in the rear.

Lack of Equilibrium

It will be agreed that lack of equilibrium, resulting from unfavorable particle size or irregular packing of the stationary phase, also from excessive rates of flow, will operate to increase the diffuseness of boundaries in a chromatogram. Martin and Synge[2] have adapted the theory of the fractionating column to chromatography, in an attempt to take explicit account of lack of equilibrium and of diffusion. This is done by dividing the column into a number of theoretical plates, of such height (H.E.T.P.) that the solution leaving a plate may be considered to be in equilibrium with the mean concentration of solute absorbed in that plate. They consider only the case of a linear isotherm and conclude that the solute initially present in one plate is, after passage through r plates, distributed, in a normal distribution, as a band with an approximate width (measured at 1/e of the maximum concentration) of $2\sqrt{r}$ plates.

The authors calculate that the H.E.T.P. in their experiments were only a few thousandths of a centimeter. This suggests that diffusion and lack of equilibrium were unimportant determinants of the conditions of the boundaries in their chromatograms. Martin and Synge derive an expression for the retardation of the point of maximum concentration of a band, which may readily be transformed to an equation identical with EQUATION 2.

The theory has not been extended to non-linear isotherms. It is apparent, however, that the effect would be to skew the distribution curve, in a manner corresponding to the changes in concentration gradients in the boundaries which are required by the equilibrium theory.

Experimental Elution Curves

Weil-Malherbe[7] has published a careful analysis of the relation of the concentration of solute to the volume of eluate, in a series of experiments with a single solute on columns of alumina and silica gel. The effects of the shape of the column, the quantity of adsorbent, the particle size, the quantity and concentration of solute, and the quantity of developing solvent are all examined. The general conclusion is that the equilibrium theory is qualitatively in agreement with the observations, but that there are some quantitative discrepancies which may plausibly be attributed to a finite rate of absorption, hydrodynamic effects, etc.

Many examples of elution curves from charcoal columns will be found in publications from the laboratory of Tiselius.[6] A few are reproduced in FIGURE 2. Curves A illustrate the fall in v_r of both boundaries, as the initial concentrations of components are increased. In all of the curves, it will be observed that the peak concentration of the leading band is greater than that of the succeeding one, and that the second boundary is somewhat more diffuse than the first. Pronounced tailing is evident in the Elution

Analysis (Curve B), also in the leading band of one of the examples (Curve C) of a Displacement Analysis.

APPLICATIONS TO MIXTURES OF AMINO ACIDS

The comprehensive analysis of a protein hydrolysate customarily takes the form of the preliminary separation of the acidic, the basic, and the neutral amino acids from one another. This is then followed by a detailed analysis of each of the three primary fractions. The preliminary step is one which depends upon the differences in electrochemical characteristics of the three types of amino acids. Two classical methods of exploiting these differences are electrodialysis and extraction with butyl alcohol. More recently, the use of columns of ion exchange agents has been suggested, as a more effective method of achieving the same results. The exchange reactions of amino acids will be discussed in a later section.

The separation from one another of amino acids of the same charge type

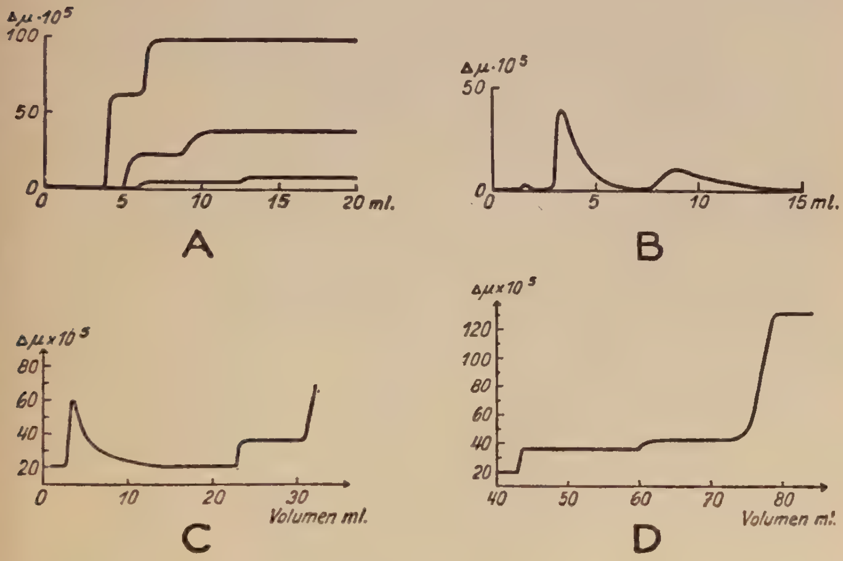


FIGURE 2.

The ordinate represents interferometer readings.

Curves A. A Front Analysis of three dilutions of a mixture of equal concentrations (0.5–0.05%) of lauric and palmitic acids in ethyl alcohol.

Curve B. An Elution Analysis of equal amounts (5 mg.) of lauric and palmitic acids with ethyl alcohol.

Curve C. A Displacement Analysis of equal amounts (10 mg.) of glucose and sucrose with 0.5% phenol (0.6 g. carboraffin).

Curve D. A Displacement Analysis of equal amounts (25 mg.) of sucrose and maltose with 0.5% phenol (1.5 g. carboraffin).

must be based upon differences in the characteristics of the side chains, other than the ionizing groups which they may contain. In the past few years, two distinctive methods of differential distribution have been developed for the separation of uncharged amino acids and peptides from one another. In one, the amino acid mixture is analyzed by passage through a column of active charcoal. In the other, a partition chromatogram is developed.

Columns of Active Charcoal

Wachtel and Cassidy[8] were the first in this country to extend chromatographic procedures to neutral amino acids. Mixtures of glycine, leucine, phenylalanine, and tyrosine were filtered through packed columns of Darco 60. Development and elution with water gave a fraction composed of glycine and leucine. The column was then further developed with 5% acetone, to separate the bands of phenylalanine and tyrosine. The column was sectioned, and the adsorbed materials in the two sections were extracted separately with 5% ethyl acetate. The mixture of glycine and leucine was separated by a second filtration through a longer column. The recoveries of glycine and leucine were quantitative. Those of the two other amino acids were 85%. Since evidence for the presence of ammonia and of other decomposition products was obtained, some oxidation on the surface of the charcoal evidently occurred.

Cheldelin and Williams[9] have measured the isotherms for the adsorption of a number of amino acids on Darco 60. From their results, we have computed the amounts of the individual amino acids which would be bound by 1 g. of Darco 60 in equilibrium with 0.1 M solutions of the pure amino acids. The results are given in TABLE 2, as amounts relative to the amount of glycine bound. It will be noted that the relative retardations in the experiments of Wachtel and Cassidy — glycine < leucine < phenylalanine < tyrosine — correspond with the relative adsorptions in equilibrated systems.

In the laboratory of Tiselius,[6] the charcoal column has been developed in a very elegant manner. The objective was to apply to the effluents and eluates from adsorption columns the methods of continuously recording concentrations, which have been so valuable in studies of diffusion and electrophoresis. The amino acid solution is slowly forced upward through the column into an optical cell placed at the end of the column. The volume flow of solution is continuously recorded, as well as the total concentration of solutes. For the latter, both refractometric and interferometric methods have been employed. The specific retardation volumes for 0.5% solutions of some 35 amino acids and peptides were first determined. A representative list is included in TABLE 2, where they may be compared with the adsorption data of Cheldelin and Williams.

The analysis of a mixture may be carried out in one of several ways. The method of Front Analysis consists of the passage of a solution until the original mixture appears in the effluent. The concentration record of

TABLE 2

RELATIVE DISTRIBUTION COEFFICIENTS AND RETARDATION VOLUMES OF AMINO ACIDS

Column I. Relative partition coefficients in butanol-water.[18]
Column II. Relative adsorption on Darco 60 at 0.1 M.[9]
Column III. Specific retardation volumes in charcoal columns (0.5% solutions).[6]
Column IV. Relative partition coefficients of the N-acetyl amino acids between chloroform and water.[12]
Column V. Relative velocities in aqueous silica gel columns from chloroform-butanol solutions.[2]
Column VI. Relative velocities in aqueous silica gel columns from cyclohexane-propanol solution.[2]

	I	II	III		IV	V	VI
Glycine	1	1	0	..	1	slow	1
Alanine	1.6	3	0	..	3.5	2	2
Amino butyric acid	3	21	..	..	15	7	..
Valine	4.5	21	3	..	20	15	10
Leucine	12.5	43	8	..	85	..	..
Isoleucine	..	43	10	..	100	30	20
Norleucine	21	43	..	..	120	..	..
Serine	..	3	..	..	0.6	..	..
Threonine	..	5	..	..	..	..	..
Methionine	..	68	..	..	27	15	4.5
Tryptophane	..	65	75	..	..	30	10
Phenylalanine	..	46	..	..	120	45	20
Proline	..	..	3	..	40	15	2
Hydroxyproline	..	..	3	..	0.6	..	..
Tyrosine	..	140	..	..	1	2	4
			pH 7	*9.9*			
Aspartic acid	..	30	..	..	..	..	..
Glutamic acid	..	..	..	..	0.6	..	..
Histidine	..	..	..	16	..	..	..
Lysine	..	27	75	10	0.6*	..	..
Arginine	..	..	40	25	..	..	..
Glycyl-glycine	..	..	3.5	..	..	slow	..
Leucyl-glycine	..	..	18	..	..	2	20
Leucyl-glycyl-glycine	..	..	30	..	..	..	..
Leucyl-leucine	..	..	..	..	..	100	..

* (Di-acetyl).

the effluent shows a series of discontinuities, alternating with plateaux (FIGURE 2). The retardation volume corresponding with each discontinuity is measured and serves to identify the component of the fraction. The increments in concentration represented by the successive plateaux are related to the original concentrations, but, as we have seen, are not, in general, identical with them.

Analyses may also be made by development and elution of bands and by displacement. Some of the characteristics of these analyses have been considered in the discussion of FIGURE 2.

Little work with neutral amino acids has been done with columns other than those of activated charcoal. Schramm and Primosigh[10] succeeded in separating tryptophane, tyrosine, and phenylalanine in columns of alumina; Karrer, Keller, and Szonyi,[11] on the other hand, have used basic zinc carbonate, to fractionate mixtures of the N-p-phenyl-azobenzene derivatives of some neutral amino acids dissolved in 95% ligroin–5% benzene. These derivatives are colored, so that the development of the chromatogram may be followed visually. The zones, in order, from the top of the column, are given as glycine, alanine, leucine, and valine.

Liquid-Liquid Distribution Columns

The work of Gordon, Martin, and Synge originated in the attempt to fractionate the N-acetyl derivatives of the amino acids, by partition between chloroform and water. Synge[12] measured the distribution coefficients and found that they covered a wide range of values. The values relative to acetyl glycine are recorded in TABLE 2, where it is interesting to compare the distribution, between chloroform and water, of the acetyl-derivatives with the distribution, between butyl alcohol and water and between charcoal and water, of the free amino acids. Synge[13] then developed a method for the complete acetylation of the amino groups, in mixtures of amino acids. With this accomplished, the analysis of the mixture of acetylated derivatives, by countercurrent distribution, was undertaken. The equipment first used was a rather cumbrous 40-chamber assembly, but partial analyses of the hydrolysis products of wool and of gelatin were satisfactorily completed.[14] Later, the process was reduced to a simple column filtration by the use of silica gel to hold the aqueous phase. An acid-base indicator was added to the aqueous phase and served to delineate the bands of the absorbed acids. During the filtration, the volume flow and the movement of the bands were recorded. As each band of acids was ejected into the effluent, it was collected separately and titrated. The component was identified by its retardation rate.

A single fractionation did not serve to separate some of the acetyl amino acids from one another. By repeating the fractionation, however, with propanol-cyclohexane as mobile solvent, additional separations were achieved. The authors claim to be able to obtain the following fractions from a mixture: phenylalanine, tyrosine, methionine, proline, the mixed leucines, and glycine plus alanine. Extensive analyses of gramicidin[15] and of tyrocidin[16] have been reported.

In the method which has been described, the N-acetyl derivatives of the hydroxyamino acids remain in the water phase. The authors propose a method of separating the hydroxyamino acids as a group in the form of their diacetyl derivatives. It has also been suggested that lysine might be conveniently separated from the mixed base fraction as diacetyl lysine. The authors have failed to separate aspartic acid from glutamic acid by their method. They have considered, but do not appear to have developed, the fractionation of the methyl esters of the acetyl amino acids by distribution between ethyl acetate and water.

Wieland and Fremerey[17] have employed the silica gel column to fractionate the copper complexes of the amino acids, using a mixture of phenol and chloroform as the perfusing solvent. The order of decreasing retardation of the amino acid complexes is given as serine and glycine — hydroxyproline — alanine — valine and leucine — proline and methionine. In view of the ability of copper to form mixed complexes with the amino acids, one must conclude that there is a continual readjustment of the composition of the complexes in the developing bands. This complication would not appear to be one which would favor fractionation.

Gordon, Martin, and Synge[2] have emphasized certain technical difficulties associated with variations in the behavior of different samples of silica gel, with the purification of solvents and with the choice of a satisfactory indicator. Others have noted the difficulty of carrying out the preliminary acetylation quantitatively. Meanwhile, the authors have sought to extend the method to the separation of the amino acids themselves, by means of the solvent pair, butanol-water. The partition coefficients for this pair had already been measured by England and Cohn[18] (TABLE 2). They show a wide range of values. Gordon *et al.* have found silica gel to be an unsatisfactory matrix for the immobile phase in this case, but report more promising results with cellulose acetate fiber.[19] The use of starch grains has also been proposed.[20]

An interesting development, from the same laboratory, has been the investigation of two-dimensional chromatograms on filter paper impregnated with water and developed with various solvents.[21] The method recalls the process of capillary analysis, described by Brown.[22] Chromatograms may be developed with as little as 1 mg. of hydrolyzed protein. The bands are visualized by spraying the paper with ninhydrin, and identified by their relative positions, using chromatograms of single amino acids for reference. Only qualitative significance is claimed for the method. However, its usefulness is attested by a paper which has just been published by Consdon *et al.*[23] On the basis of analyses with two-dimensional chromatograms, they question the presence of norleucine in the proteins of ox brain. Since the latter has been the chief source of natural norleucine, the result is important. An interesting part of their paper deals with the identification of Thudichum's original preparations of "glycoleucine" with *dl*-leucine. It was Abderhalden's conclusion that glycoleucine was norleucine, that first led to the inclusion of norleucine in the list of naturally-occurring amino acids.

Separations Based Upon Ion Exchange

Between pH 3 and 6, the basic amino acids are present in solution as cations, the dicarboxylic amino acids as anions, and the neutral amino acids as dipolar ions. A proper use of ion exchange reagents in this region of pH should make possible a quantitative separation of the three classes of amino acids from one another.

Acidic Amino Acid Exchange

Although it has long been recognized that a variety of materials (*e.g.*, certain activated clays and the gels of various metallic hydroxides) exhibits the capacity to exchange bound anions, only recently has any attempt been made to exploit this property in the analysis of amino acids. Successful separations of the dicarboxylic amino acids have been reported with titania gel[24] and with acid-washed alumina.[25] In the latter case, solutions of the sodium salts of the amino acids in water or in 80% alcohol were treated with the reagent. Washing the adsorbent with acetate buffer served to elute the glutamic acid and subsequent treatment with NaOH recovered the aspartic acid. This useful separation is not easily explained on a purely electrochemical basis. The isoelectric points of the two dicarboxylic amino acids are so close together that it is difficult to understand why their adsorbed ions should behave so differently toward an acetate buffer.

A more promising type of reagent is that represented by the polyamine-formaldehyde resins which have become commercially available, within the past few years. These, with the complementary cation-exchange resins prepared from phenols and formaldehyde, have been successfully employed to remove electrolytes from a variety of natural and industrial waters. Their availability has naturally stimulated interest in their use as analytical reagents. Some aspects of the behavior of typical representatives of the anion exchange resins toward amino acids have been investigated by Turba, Richter, and Kuchar,[26] by Englis and Fiess,[27] and by Cleaver, Hardy, and Cassidy.[28] A method of using one of them in the fractionation of protein hydrolysates has been developed,[29] and the dicarboxylic amino acids in three standard crystalline proteins have been estimated by this method with satisfactory results.

These resins merit further investigation. They are quite stable in alkaline and reasonably so in acid solutions. They have a high exchange capacity and can be used, either in the form in which their functional groups are present as free basic groups (ROH), or in the form in which they are in salt union with an anion (*e.g.*, RCl). The latter will exchange Cl^- for other anions, whereas ROH may be used to remove free acids from solution — *i.e.*, exchange OH^- for the anions of these acids. This capacity is, however, limited by the nature of the assembly of ions in the solution. The reaction, $ROH + A^- + B^+ \longrightarrow RA + BOH$, can move to the right only to the extent that ROH is a stronger base than BOH. Now, the functional groups of the resins are weak bases, in the sense that they form salts only on the acid side of neutrality. On the other hand, the cations present in the hydrolysate of a protein are converted to their free bases only in the region of pH 9. When, therefore, the hydrolysis products of a protein are filtered through a bed of resin of the type under consideration, the amount of acid removed is limited to that which is in excess of the bases ionized at neutrality. Quantitative removal of any particular anion can ultimately be achieved only by reacidifying the filtrate and refiltering several times through fresh columns of ROH. This tedious procedure could obviously be avoided if: (a) the bases were removed before the attempt to remove the acids; (b) a resin were available whose basic groups were strong enough to raise the pH of the solution above 9. The former method was adopted by Cannan.[29] The latter should also be investigated.

Exchange of Basic Amino Acids

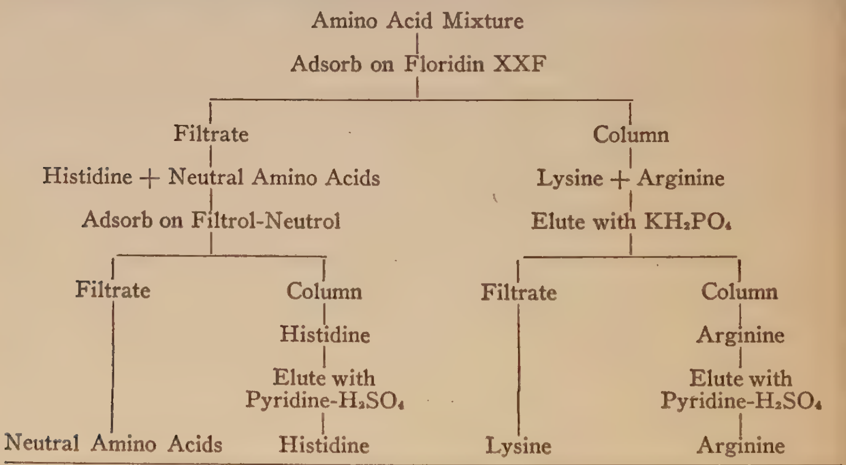
Alumina has amphoteric properties. Reference has already been made to its ability to bind anions after activation by acid. The untreated gel, on the other hand, has, according to Wieland,[30] base exchange capacity for the amino acids. If the pH of the solution is as high as 8, only lysine and arginine are adsorbed. From acid solutions, all of the bases are removed together, with some non-basic amino acids. It is stated, however, that the latter may readily be eluted with a pyridine-sulfuric acid buffer. This chain of events is entirely consistent with the view that the reactions involved are essentially processes of ion exchange.

Lloyd's reagent[31] has been used to adsorb the basic amino acids, and extensive observations[32] have been made with two bleaching earths, Filtrol-Neutrol and Floridin XXF Extra. The former was found to remove all of the bases from a neutral solution of the amino acids, whereas Floridin XXF only removed lysine and arginine. This difference may be attributed to the weak basic properties of histidine. If so, it may be inferred that Floridin XXF is distinctly less acidic than Filtrol-Neutrol.

Further developments in the work with these two agents are indicated by the scheme for the complete fractionation of the bases,[33] shown in TABLE 3.

TABLE 3

ANALYSIS OF THE BASIC AMINO ACID FRACTION AFTER WIELAND[33]



The methods adopted for eluting the bases seem to depend on exchange of the potassium or the pyridinium ion for the amino acid cation in weakly acid solution. Why lysine and arginine should behave so differently toward KH_2PO_4, finds no electrochemical explanation.

An interesting development in the use of Filtrol-Neutrol was made by Waldschmidt-Leitz and Turba.[34] With the aid of this adsorbent, they succeeded in separating the partial hydrolysis products of a tryptic digest of clupein into three fractions, with ratios of total N to amino N of 5, 9, and 17, respectively. Evidently, a substantial fractionation on the basis of the arginine contents of the peptides was achieved.

It should be recorded that neither alumina nor the bleaching earths can be regarded as efficient exchange agents. Their exchange capacities appear to be of the order of 0.1 meq. per g. Turba also notes that it is necessary to work rapidly, if quantitative recoveries are to be obtained. Finally, the appearance of phosphate, sulfate, and pyridine in the final fractions detracts from the elegance of the proposed analytical procedures.

Many years ago, Whitehorn[35] used permutit to isolate the basic amino acids from a mixture. More recently, the phenol-formaldehyde resins have been employed. Those which have found particular favor as cation exchange agents are those which have been sulfonated to increase their acidity. They are stable to quite concentrated solutions of mineral acids and, consequently, can be prepared and used in the form of their free acids (RH). Toward solutions of electrolytes, indeed, RH behaves as a strong acid, showing a measurable ability to exchange H^+ for other cations in solutions in which H^+ is as great as 1M. This strong tendency of the sulfonated cation-exchange resins to exchange H^+ for other cations is in notable contrast with the anion-exchange resins, whose affinity for OH^- is so much greater than their affinity for other anions. The strongly acid nature of RH has advantages and disadvantages, from the point of view of amino acid analysis. A notable advantage is that amino acids may be removed from acid solutions by RH, without adding to the solution any analytically undesirable cation. On the other hand, due to the acid strength of RH, it has been found necessary to use high concentrations and large volumes of mineral acids to effect a quantitative elution of organic bases from the resin.[36]

Block[37] has reported success in the use of RH for the removal of the basic amino acids from mixtures of amino acids previously brought to pH 6 by treatment with an acid binding resin (ROH). The column of RH containing the bases was first washed with water and then eluted with 2 M HCl. Under these conditions of operation, it is probable that there is also extensive adsorption of neutral amino acids. From our experience, it is difficult to remove these simply by washing with water. Freudenberg *et al.*[38] have acknowledged this. They formed their chromatogram with a solution of the hydrochlorides of the amino acids. They then removed the neutral amino acids, which were adsorbed by washing with a pyridine solution before elution of the bases with acid. It is interesting to observe that pyridine was used, in this case, to remove the neutral amino acids by suppressing their ionization. In the experiments of Turba,[32] it was employed to displace the basic amino acids by exchange. It is evident that the conditions of operation must be carefully controlled to suit the particular objective.

By reason of high capacity, uniformity of product, stability, and long life, the ion exchange resins are attractive analytical agents. The experiments quoted above suggest that they may have some value in the analysis of mixtures of amino acids. An ideal scheme for their use in the separation of amino acids into groups of the same charge type would be somewhat as follows: A basic resin (ROH) would be added to the hydrolysate of a protein to bring the pH close to the isoelectric point of the more basic amino acids (pH 9–10). This would remove the dicarboxylic acids (and, perhaps, small amounts of neutral amino acids), together with the acid which had been used to effect hydrolysis. The filtrate would then be treated with an acidic resin (RH) to reduce the pH to about 4 to 6, thereby removing the basic amino acids and ammonia. It might be possible, indeed, to conduct this step in two stages. Addition of RH to pH 8 should remove arginine and lysine and a further addition to pH 5 should absorb the less strongly basic histidine. The solution which remained would contain the neutral amino acids uncontaminated by foreign reagents and could be directly submitted to chromatographic analysis in charcoal or other columns. The basic and the acidic amino acids could be recovered from the resins as hydrochlorides by elution with hydrochloric acid. This attractive scheme requires an acidic resin with weaker acid properties than the sulfonic acid resins and a basic resin with stronger basic properties than those presently in use. In time, such reagents should become available.

Since the advantages of the hydrogen cycle cannot, as yet, be fully exploited in the separation of the amino acids, an alternative is to employ columns of resin salts and remove the acidic and the basic amino acids by exchange in neutral solution. For example, the hydrolysate might be neutralized to about pH 5 by addition of a commercial basic resin (ROH) with partial adsorption of the dicarboxylic amino acids. Filtration of the solution through a column of RCl would then remove the remainder of these acids by exchange with chloride. Finally, the basic amino acids could be exchanged for ammonium or other cation in a column of the appropriate salt of RH. The disadvantage of this procedure is that the final filtrate and the acid eluate containing the basic amino acids would both be contaminated with much chloride of the exchange base.

A possibility which has not, hitherto, been systematically explored, is that the affinities of a series of amino acids of the same charge type for an ion exchange agent may differ sufficiently from one another to make possible their chromatographic separation. The separation of arginine from lysine and of glutamic acid from aspartic acid, which has been cited above, may depend on such differences. There is some evidence, also, that the neutral amino acid hydrochlorides may differ in their exchange affinities for acidic resins.[26] As a contribution to this problem, we have examined the equilibrium conditions in the reaction of two commercial sulfonic acid resins with a series of uni- and divalent cations. Both of the resins exchanged rapidly and essentially reversibly and gave substantially the same equilibria with a given cation. For any one sample of resin, the exchange capacity was independent of the nature of the ion, but the exchange constants showed significant differences.

The exchange equilibria for a univalent ion were satisfied by the isotherm:

$$RB = \frac{T_R.K_1b^p}{h^p + K_1b^p}, \tag{7}$$

and those for a divalent ion, by:

$$2RB = \frac{T_R.K_2b^p}{h^{2p} + K_2b^p}. \tag{8}$$

The quantities in these equations are defined as follows:

T_R — the equivalent capacity of 1 gm. resin.
RB — the moles of cation bound by 1 gm. resin.
h — the equilibrium concentration of H^+ in solution.
b — the equilibrium concentration of cation in solution.
K, p — constants for the cation.

A representative series of experimental values of K and p are given in TABLE 4. It will be seen that K is, in general, greater than unity, and p is less. In a series of related ions, K increases and p diminishes as the mass of the cation increases. Amongst the organic cations, however, specific structural influences are also evident. In the case of the amino acids, the differences in K are less significant than are the differences in p.

In order to evaluate the significance of these differences to the progress of exchange reactions in columns, we have extended the equilibrium theory of chromatography to reactions which may be represented by equations of the type of 7 and 8. Space does not permit the development of the

TABLE 4

EXCHANGE CONSTANTS FOR AMBERLITE IR100

	K_1	p		K_2	p
Lithium	0.65	1.05	Magnesium	1.9	0.52
Sodium	1.25	1.00	Calcium	3.1	0.50
Potassium	2.1	0.90	Strontium	3.5	0.47
Rubidium	3.0	0.85	Barium	4.4	0.47
Caesium	3.5	0.82			
Ammonia	2.0	0.91			
Methylamine	2.7	0.85			
Ethylamine	3.4	0.87			
Butylamine	8.1	0.87			
Diethylamine	5.4	0.77			
Tetraethylammonium	2.1	0.58			
Alanine	1.25	0.90			
Leucine	1.8	0.71			
Tyrosine	2.2	0.63			
Methionine	2.0	0.75			
Glutamic Acid	0.4	1.10			
Lysine	2.2	0.91		5.0	0.45
Histidine	4.0	0.80		5.6	0.42
Arginine	6.0	0.75		7.2	0.35

argument in this review, but a brief statement of some qualitative deductions from it is pertinent to the immediate question.

In the development by acid of the band of a single solute, the sharpness of the boundaries will depend primarily on the value of K. If K is greater than unity — as seems to be the general case for sulfonic acid resins — the leading boundary will be sharp and the trailing boundary will be diffuse. In other words, a boundary will be sharp if it is one in which an ion of higher affinity for the resin is replacing one of lower affinity. The greater the value of K, the greater will be the attenuation of the tail of the band during development. This spreading out of the tail will be further accentuated if p is less than unity. Consequently, a cation for which K is large and p is small will rapidly spread out, during development, to form a wide, dilute band. It will be recovered in the eluate in low concentration, and large volumes of acid will be required to remove it completely from the column. This conforms with experience which has shown that it is difficult to obtain quantitative recovery of complex organic univalent cations by elution with concentrated solutions of mineral acids. The same difficulty has been encountered in the elution of the basic amino acids.

The theory indicates that conditions can be chosen in which the retardation rate of a sharp boundary becomes a function of K and of p. Under these conditions, the retardation is increased by a rise in K and by a fall in p. This would suggest that bands of cations with different constants can be caused to move at different rates. These differences in rates, however, are small, compared with the simultaneous differences in the extent of spreading of the band.

It is not possible to make any general statement about the progress of a mixed band as a function of the constants of the components. It is to be expected that the component of higher K or lower p will progressively displace the other component, until two contiguous bands are formed. While this is in progress, however, the spreading of the band under the diluting influence of the acid which is being used for development will operate against an effective separation of the two components.

The equilibrium theory leads to the conclusion that the effective method of preventing the tailing of bands during development is to employ, as the developing agent, a cation with a greater affinity for the resin than any of the cations in the band. The affinity of the hydrogen ion in the presence of a particular cation is given by the reciprocal of the K of that cation. For sulfonic acid resins, this will, in general, be less than unity. Consequently, acids are unfavorable developing agents for cations bound to sulfonic acid resins. Instead of producing a chromatogram with sharply defined bands of high concentration, the acid tends to diffuse into the bands diluting them and spreading them out. One is led to the conclusion that more satisfactory chromatographic results would be obtained by using as the developing agent the salt of a cation of high exchange affinity. Salts of aniline, pyridine, or similar weak bases would seem to be the most promising ones, since they can be easily removed from the eluates. An alternative which remains to be explored is the use of an exchange agent whose functional groups are less strongly acidic than is the sulfonic acid grouping.

BIBLIOGRAPHY

1. **Craig, L. C.**
   1944. J. Biol. Chem. **155**: 519.
2. **Martin, A. J. P., & R. L. M. Synge**
   1941. Biochem. J. **35**: 1358.
   Gordon, A. H., A. J. P. Martin, & R. L. M. Synge
   1943. Biochem. J. **37**: 79.
   Gordon, A. H., A. J. P. Martin, & R. L. M. Synge
   1944. Biochem. J. **38**: 65.
3. **Wilson, N.**
   1940. J. Am. Chem. Soc. **62**: 1583.
4. **De Vault, D.**
   1943. J. Am. Chem. Soc. **65**: 532.
5. **Weiss, J.**
   1943. J. Chem. Soc. 297.

6. **Tiselius, A.**
1943. Kolloid Zeit. **105**: 101. (Bibliography cited.)
1942. Advances Colloid Sci. **1**: 81.
1944. Several papers in The Svedberg Volume. Almquist and Wiksells. Upsala.
7. **Weil-Malherbe, H.**
1943. J. Chem. Soc. 303.
8. **Wachtel, J. L., & H. G. Cassidy**
1943. J. Am. Chem. Soc. **65**: 665.
9. **Cheldelin, V. H., & R. J. Williams**
1942. J. Am. Chem. Soc. **64**: 1513.
10. **Schramm, G., & J. Primosigh**
1943. Ber. chem. Ges. **76B**: 373.
11. **Karrer, P., R. Keller, & G. Szonyi**
1943. Helv. Chim. Acta. **26**: 38.
12. **Synge, R. L. M.**
1939. Biochem. J. **33**: 1913.
13. **Synge, R. L. M.**
1939. Biochem. J. **33**: 1918.
14. **Martin, A. J. P., & R. L. M. Synge**
1943. Biochem. J. **35**: 91.
15. **Gordon, A. H., A. J. P. Martin, & R. L. M. Synge**
1943. Biochem J. **37**: 86.
16. **Gordon, A. H., A. J. P. Martin, & R. L. M. Synge**
1943. Biochem. J. **37**: 313.
17. **Wieland, T., & H. Fremerey**
1944. Ber. chem. Ges. **77B**: 234.
18. **England, A., Jr., & E. J. Cohn**
1935. J. Am. Chem. Soc. **57**: 634.
19. **Gordon, A. H., A. J. P. Martin, & R. L. M. Synge**
1943. Biochem. J. **37**: xiii.
20. **Elsdon, S. R., & R. L. M. Synge**
1944. Biochem. J. **38**: ix.
21. **Consdon, R., A. H. Gordon, & A. J. P. Martin**
1944. Biochem. J. **38**: 224.
22. **Brown, W. G.**
1939. Nature **143**: 377.
23. **Consdon, R., A. H. Gordon, A. J. P. Martin, O. Rosenheim, & R. L. M. Synge**
1945. Biochem. J. **39**: 251.
24. **Johnson, M. B.**
1938. Dissert. Catholic Univ. Am.
25. **Turba, F., & M. Richter**
1942. Ber. chem. Ges. **75B**: 340.
Wieland, T.
1942. Ber. chem. Ges. **75B**: 1001.
Wieland, T.
1942. Naturwissenschaften **30**: 374.
26. **Turba, F., M. Richter, & F. Kuchar**
1943. Naturwissenschaften **31**: 508.
27. **Englis, D. T., & H. A. Fiess**
1944. J. Ind. Eng. Chem. **36**: 604.
28. **Cleaver, C. S., R. A. Hardy, Jr., & H. G. Cassidy**
1945. J. Am. Chem. Soc. **67**: 1343.

29. **Cannan, R. K.**
1944. J. Biol. Chem. **152**: 401.

30. **Wieland, T.**
1942. Z. physiol. Chem. **273**: 24.

31. **Ackerman, D., & H. Fuchs**
1936. Z. physiol. Chem. **240**: 198.

32. **Turba, F.**
1944. Ber. chem. Ges. **74B**: 1829.

33. **Wieland, T.**
1943. Die Chemie **56**: 213.

34. **Waldschmidt-Leitz, E., & F. Turba**
1940. J. Prakt. Chem. **156**: 55.
1941. Ibid. **158**: 72.

35. **Whitehorn, J. C.**
1923. J. Biol. Chem. **56**: 751.

36. **Herr, D. S.**
1945. J. Ind. Eng. Chem. **37**: 631.

37. **Block, R. J.**
1942. Proc. Soc. Exp. Biol. & Med. **51**: 252.

38. **Freudenberg, K., H. Walch, & H. Molter**
1942. Naturwissenschaften **30**: 87.

MICROBIOLOGICAL METHODS IN AMINO ACID ANALYSIS

By Esmond E. Snell

Department of Biochemistry, University of Wisconsin, Madison, Wisconsin

INTRODUCTION

For many years, investigators interested in microbial nutrition have utilized the growth responses of microorganisms to detect the presence of substances favorable or necessary for growth of such organisms on laboratory media. In many instances, such responses could be controlled to yield results sufficiently quantitative to permit isolation or identification of the active substances involved. As a result of such investigations, the nutritional requirements of a large number of microorganisms are now known. Once the growth requirements of an organism are known, it becomes possible to devise media which are deficient in a single substance only and to employ the response of the organism to additions of that substance, for the quantitative estimation of the substance in natural materials. Such procedures depend upon the observed fact that, in the presence of limiting amounts of a substance essential for growth, the amount of growth obtained (or the amount of some metabolic product used to follow growth) is a function of the concentration of essential nutrilite and increases to a maximum as the concentration of the nutrilite is increased. A curve relating the amount of growth to the concentration of the pure essential substance is thus obtained, and can be used to assess the amount of the essential substance present in mixtures of varying complexity. Such procedures have met with widespread success in the estimation of many vitamins. Similar techniques are now being successfully applied to the estimation of the amino acids.

While any organism which specifically requires a given amino acid for growth can, theoretically, be used in this manner for its determination, only a few have so far proved useful. Foremost among these are various species of lactic acid bacteria. These organisms possess a number of characteristics which make them especially valuable for this type of work. Their nutrition is complex. A variety of vitamins and amino acids is required, among other substances, for growth. Extensive study has elucidated these requirements, so that it is possible to grow many of these organisms on synthetic media. Media deficient in any one of the numerous substances required by them can thus be readily prepared and studied.[1] The organisms grow rapidly, and their growth can be followed readily, either turbidimetrically, or by titrating the acid produced as a metabolic product. Also of value is the extensive experience concerning the behavior of organisms of this group under varied circumstances which has accumulated since their introduction as assay agents for the vitamins.[2, 3]

Organisms not belonging to the group of lactic acid bacteria, which have been used or recommended for assay of one or more amino acids, include *Eberthella typhosa,*[4] and mutant cultures of *Neurospora crassa*[5, 6] and *Escherichia coli.*[7] This list will probably be greatly extended, as investigation in this field continues.

Only methods which involve active growth of microorganisms have been mentioned above. A number of other methods or potential methods for various amino acids involve the use of microorganisms, but are more properly classified as enzymatic methods. These have their prototype in the arginase method for arginine,[8] and will be mentioned only in passing. Resting or killed cells of microorganisms are here used as a convenient source of specific enzymes. An example is the method for determination of *l*-lysine, which utilizes a specific lysine decarboxylase from *Bacterium cadaveris.*[9] Decarboxylases are also available from other microorganisms for tyrosine, histidine, arginine, ornithine,[10] and glutamic acid,[11] and may be found for additional amino acids, as well.

The discussion below will consider, primarily, certain aspects of the use of lactic acid bacteria as assay agents. A more extended review is available elsewhere.[1]

THE ASSAY METHODS

A semi-synthetic basal medium which will support growth of a number of lactic acid bacteria is shown in TABLE 1. Omission of tryptophane produces a medium no longer capable of supporting growth of organisms which require tryptophane. A very similar medium has been recommended for the determination of tryptophane.[15] The hydrolyzed casein present in such a medium can be replaced by a mixture of the twenty or so commonly occurring amino acids (5 to 20 mg. of each per 100 cc.), without influencing adversely its ability to support growth. When this is done, the resulting medium closely resembles those presently used for the analysis of amino acids. With media of such complexity, an almost infinite number of variations in concentrations of individual ingredients is possible, so that a number of very closely related media are used by different investigators in this field. It should not be inferred that such variations are always unimportant, although they appear to be so in some instances. In some cases, they may make the difference between securing excellent growth and acid production of a test organism, with highly reproducible results, and securing comparatively poor growth or acid production, with results which have little utility. Because of their similarity and the general lack of comparative data secured with closely related media, there is little merit in comparing the composition of such media in further detail.

TABLE 1

A SEMI-SYNTHETIC BASAL MEDIUM FOR LACTIC ACID BACTERIA*

	Amount per 100 cc. of medium
Casein hydrolysate, charcoal treated (or amino acids)	0.5 gm.
Sodium acetate	0.6 gm.
Glucose	1.0 gm.
Cystine	10.0 mg.
Tryptophane	10.0 mg.
Uracil	1.0 mg.
Adenine	1.0 mg.
Guanine	1.0 mg.
Riboflavin	10.0 γ
Calcium pantothenate	10.0 γ
Nicotinic acid	10.0 γ
Pyridoxine†	100.0 γ
Biotin	0.04 γ
Folic acid (*L. casei* factor, vitamin B_c)	0.1 γ
Thiamine	10.0 γ
p-Aminobenzoic acid	10.0 γ
Inorganic salts‡	

* This medium, with pyridoxine omitted, was used for assay of pseudopyridoxine;[12] it represents a slight modification of an earlier medium used for assay of nicotinic acid.[8] The medium can be made to support growth of additional organisms of the lactic group, by addition of asparagine and glutamine.[13]

† Pyridoxamine or pyridoxal can be substituted in much smaller amounts, with as good or better results.[14]

‡ These include, per 100 cc. of medium, 50 mg. of K_2HPO_4, 50 mg. KH_2PO_4, 20 mg. $MgSO_4.7H_2O$, 1 mg. NaCl, 1 mg. $FeSO_4.7H_2O$ and 1 mg. $MnSO_4.4H_2O$.

Once a medium supporting adequate growth of the test organism is available, the amino acid requirements of the organism can be determined by omitting in turn each of the amino acids present in the medium, and determining the effect of such omission on growth and acid production. In this way, the data in TABLE 2 were secured.

Eighteen of the commonly occurring amino acids are required for one or another of the organisms so far investigated. It can be expected that application of this knowledge will eventually make possible the microbiological determination of, at least, these eighteen amino acids. It is also

TABLE 2

AMINO ACID REQUIREMENTS OF VARIOUS LACTIC ACID BACTERIA*

Organism and Reference	Aspartic acid	Glutamic acid	Arginine	Histidine	Lysine	Alanine	Cystine	Glycine	Isoleucine	Leucine	Methionine	Phenylalanine	Proline	Serine	Threonine	Tryptophane	Tyrosine	Valine
Lactobacillus arabinosus[17-19]	±	+	±	—	±	±	+	—	+	+	±	±	—	—	±	+	±	+
Lactobacillus casei[20]	+	+	+	±	±	±	+	—	±	+	±	+	—	+	±	+	+	+
Lactobacillus delbrueckii[21]	+	+	+	—	±	±	+	—	+	+	±	+	—	+	±	+	+	+
Streptococcus fecalis[22, 23]	+	+	+	—	+	+	±	+	+	+	—	—	—	+	+	+	±	+
Streptococcus lactis[24]	—†	—†	+	±	±	—	—	—	+	+	+	±	—	±	—	—	—	+
Leuconostoc mesenteroides[25]	+	+	+	+	+	±	+	+	+	+	+	+	+	+	+	+	+	+

* Norvaline, norleucine, and hydroxyproline have not been found essential for any organism yet investigated. + indicates the amino acid is essential; ± that some growth, but not maximum growth, occurs in the absence of the amino acid; and — that the amino acid is non-essential.

† Glutamine and asparagine were essential for growth and undoubtedly supplied the organism with these amino acids.

evident that several organisms might be used for the determination of various single amino acids.

Data such as those listed in TABLE 2 should be accepted with certain reservations. Individual cases may be found where they are in error, due to impurities in the amino acids used.[23, 26] The data should be taken as representative of the requirements of a species only, since it is well known that the requirements of different strains or cultures of a single species may vary considerably. Furthermore, a single culture may, under special circumstances, lose its requirement for an amino acid, as has recently been shown for *Lactobacillus arabinosus,* in the case of tryptophane.[27] It has also been shown repeatedly that the requirement of an organism for a given amino acid is, in some cases, dependent upon the composition of the medium. Only rarely is it possible to culture an organism on a mixture of amino acids which contains only those shown to be essential when omitted individually from a complete mixture. Almost invariably, additional amino acids are required for growth under such conditions.[18, 24] Stokes and Gunness[21] found that *Lactobacillus delbrueckii, L. casei,* and *L. arabinosus* required lysine, threonine, and alanine, if grown in a medium containing pyridoxine. If an equal amount of pyridoxamine or pyridoxal[28] were used, these amino acids were not required. However, the requirement of *Streptococcus fecalis* for these amino acids was unaltered by the presence of pyridoxamine.

It thus appears that media containing a complete assortment of amino acids and vitamins are most reliable for use in assay. Such complete media also permit one to avoid, to a large extent, the inhibitory action of individual amino acids, which is often apparent when incomplete amino acid mixtures are used.[20, 29, 30]

Except for this generalization, what principles should guide one in the selection of an assay medium or method of assay? A number of supposed requirements which should be fulfilled by a microbiological assay have been suggested from time to time. Some of these are listed below:

(a) The basal medium should support little or no growth, in the absence of the substance under investigation, but should support as good growth when this substance is added as can be obtained with tissue extracts.[31]

(b) The basal medium should contain the minimum concentration of all the nutrients required to produce a standard curve of maximal slope.

(c) The standard curve should be reproducible within close limits and approximately linear over the segment to be employed for analytical purposes.

(d) The substance assayed must be essential for the growth of the organism.

(e) The assay should be applicable over a four-fold, or even greater, range of concentration.[32]

That these characteristics are desirable in an assay, when they can be obtained, is probably beyond question. The author does not believe, however, that they are either necessary, or sufficient, to insure that an assay method will give reliable results. Rigorous application of (a) would exclude all of the microbiological methods for amino acids, which have so far appeared, which utilize lactic acid bacteria, as well as a widely used method for nicotinic acid.[2, 23] In none of these methods does growth of the test organism proceed as far, or as rapidly, as when large amounts of tissue extracts are added. Despite this fact, figures of proved reliability can be obtained with such media. The fact that faster and better growth can be achieved when tissue extracts are added to media presently available, indicates that, as knowledge of bacterial nutrition develops, it should be possible to improve these media, so far as growth is concerned, and to shorten the time required for assay. Whether such improved media will permit assays of increased reliability and precision, remains to be seen. Recent observations[34] showed that, in the case of an arginine-requiring mutant of *Escherichia coli,* a standard curve of maximal slope was obtained on a medium which yielded clearly erroneous figures for the arginine content of various samples. Suitable supplementation of the medium decreased the slope of the standard curve and permitted acceptable figures for arginine to be secured. Thus, (b) is not always applicable. Two of the better methods presently available[35, 36] for the determination of vitamin B_6 use organisms which grow extensively, under some conditions, when vitamin B_6 is not present. Thus, (d) is not generally applicable. To set up such arbitrary requirements, at the present stage of development of microbiological assays, seems unjustified. Fulfilment of two conditions should suffice to gain acceptance for a microbioassay: (a) the method should be workable, wherever tried by a capable operator, and (b) it should give a correct answer.

Where figures of proved reliability are available for the amino acid content of a protein, it is easy to determine whether a microbiological method is giving the correct answer for that protein. Such figures, however, are altogether too few. Furthermore, even though a method can be shown to give reliable figures on one or more protein samples, one cannot conclude that it is generally applicable to samples of widely varying composition, especially when such samples contain matter of non-protein origin. In the absence of analytical figures of proved accuracy for most materials, other criteria for determining the reliability of a proposed method are required. These can be listed as follows:

(a) Quantitative recoveries of added amino acids must be obtainable.

(b) Identical results, within experimental error, must result from repeated assay.

(c) The amino acid to be determined should be the only naturally occurring substance present in the sample which permits growth in the deficient medium.

(d) Values calculated from various assay levels must agree.

(e) Results obtained with different test organisms should agree.

The first three of the above points are obvious and are shared by all analytical methods. The latter two points deserve special consideration.

In setting up a microbiological test, the selected medium, deficient in a single amino acid, is prepared at twice its final concentration. Standard solutions of the amino acid and of the samples are added to a series of test tubes, in amounts calculated to produce several gradations in growth of the test organisms between no growth and maximum possible growth. Contents of all tubes are then diluted to an equal volume with water, and the same volume of medium is added. After capping, all tubes are sterilized, cooled, and inoculated with a suspension of cells of the test organism, secured from an actively growing culture. All tubes are then incubated at constant temperature, usually for 72 hours. After this, the response to the amino acid is determined by titration of the acid produced during growth or by quantitative measurement of culture turbidity. A standard curve is then constructed, relating the response of the organism to the concentration of the amino acid. The amount of amino acid present in the various samples is determined by interpolation of the response in the sample tubes on to this standard curve. The final result taken is an average of values calculated from several levels on the standard curve. Statement (d), above, simply means that these values must agree, regardless of the portion of the standard curve (within the working range) from which they are taken. Failure to do so means that the organism is not responding to the sample in the same manner that it responds to the standard, but is being affected, also, by other substances present in the sample. A consistent drift in values, either up or down, as the assay levels increase, is especially dangerous.

Where the dose-response curve is linear, this is equivalent to saying that the slope of the dose-response curve for the standard should bear a constant ratio to the slope of the dose-response curve for the sample, throughout the linear portion of the curves. In the case of such linear response to the standard, Wood[37] has pointed out that plotting curves for both standard and sample, and calculating the content of the sample from a comparison of the slopes of the curves, will yield more accurate results than the traditional method of calculation, and will often reveal errors not apparent when the latter method is used. Unfortunately, in most amino acid assays so far published, the response of the organism to increasing concentrations of the standard approaches, but does not reach, linearity.

From TABLE 2, it is apparent that several organisms are often available which require a single amino acid. In many instances, the nutritional requirements of these organisms differ significantly. In such cases, use of more than a single organism for assay is very instructive. If the same assay value is obtained with two or more organisms, the probability that it is correct is greatly increased. It is unlikely that a number of different organisms would respond to interfering materials in the same manner and to the same extent. The importance of this consideration, in the field of vitamin assay, has recently been emphasized by Leonian and Lilly.[38]

TABLE 3

ARGININE CONTENT OF VARIOUS PROTEINS AS DETERMINED BY DIFFERENT METHODS

Protein	*Lactobacillus casei*[39]		*Streptococcus fecalis*[39a]		Other methods
	Per cent	Ave.	Per cent	Ave.	Per cent
Ovalbumin	5.79, 5.49	5.6	5.72, 5.79, 5.83, 5.50	5.7	5.66 (flavianate[40])
Casein	3.81, 3.61	3.71	3.82, 3.85, 3.53, 3.58, 3.60, 3.84	3.71	3.72 (" [40])
Horse hemoglobin	3.53, 3.47	3.50	2.8	2.80	3.60 (" [40]) 3.70 (isotope dilution[41])
β-Lactoglobulin	2.92, 2.90	2.91	2.98, 2.95, 2.70	2.88	2.89 (flavianate[42])
Silk fibroin	0.98, 0.96	0.97	0.99	0.99	0.95 (flavianate[43])

These considerations can best be illustrated by referring to actual data. In TABLE 3 are shown comparative data for arginine, as determined with two test organisms on slightly different media. The values obtained on consecutive assays are given, together with the averages and comparative data from the literature. Values obtained with both organisms agree very well, except in the case of horse hemoglobin. Here, it is obvious that *Streptococcus fecalis* is giving an erroneous value. In the absence of reliable data from the literature, the discordant data, in this case, would call for further investigation. The results given in the table for consecutive assays on a single hydrolysate also give some idea of the precision with which a given result can be obtained.

Similar data for lysine are given in TABLE 4. Two different test organisms give very concordant values, which agree well with the best values

TABLE 4

LYSINE CONTENT OF VARIOUS PROTEINS AS DETERMINED BY DIFFERENT METHODS

Protein	*Streptococcus fecalis*[39a]		*Leuconostoc mesenteroides*[39a]		Other methods
	Per cent	Ave.	Per cent	Ave.	Per cent
Casein	7.83, 7.39, 7.34, 7.76	7.6	7.50, 7.54, 8.0, 7.48, 7.70, 6.70	7.5	7.5–8.1 (lysine decarboxylase[9]) 7.5 (lysine decarboxylase[44]) 8.3 (microbiological[45])* 6.25 (isolation of picrate[46])
β-Lactoglobulin	10.4, 10.0, 10.4, 10.4, 10.6	10.4	10.5		11.4 (isotope dilution[41]) 10.55 (isolation[47])
Horse hemoglobin	8.7		8.45		8.6 (isotope dilution[41])
Silk fibroin	0.90, 0.90	0.90	0.58		0.60 (microbiological[45])* 0.85 (isolation[48]) 0.25 (isolation[49])

* *Leuconostoc mesenteroides* was the test organism here. The assay medium differed in composition from that used above.

in the literature, except in the case of silk fibroin, where the two organisms differ significantly. The value obtained with *Leuconostoc mesenteroides* checked well the value obtained by another investigator with the same organism, but does not check that given by *Streptococcus fecalis*. Further investigation is indicated in such a case. It should be noted that, in this case, individual determinations with either organism are quite variable, so that an average of results of several assays should be taken to secure an accurate figure. Zittle and Eldred,[9] using the lysine decarboxylase method, noted variation from 7.0 to 8.1% in the apparent lysine content of casein, depending upon the pretreatment given and the method of hydrolysis used.

Table 5

Comparative Values for Aspartic Acid and for Histidine as Determined with *Lactobacillus delbrueckii* and *Leuconostoc mesenteroides*

		Per cent found		
Protein	Amino acid	*L. delbrueckii* 5[16]	*L. mesenteroides* P-60[50]	Other methods
Casein	Aspartic acid	6.1	7.2	6.7 (isolation[51])
β-Lactoglobulin	" "	9.3	11.5	9.9 (isolation[52]) 11.3 (isotope dilution[41])
Egg albumin	" "	8.0	9.3	8.1 (isolation[52])
Horse hemoglobin	" "	..	10.3	10.3 (isotope dilution[41]) 6.8 (isolation[51])
		L. delbrueckii 3[39a]	*L. mesenteroides* P-60[39a]	
Casein	Histidine	3.2	3.2*	2.5 ± 0.3[53]
Gelatin	"	0.84	0.83	0.40[53]

* Since this manuscript was written, a paper by Dunn, *et al.* (J. Biol. Chem. 159: 653. 1945) describing the microbiological determination of histidine with *L. mesenteroides* has appeared. They found 3.1% of histidine in casein, in excellent agreement with the above figure.

With aspartic acid, values obtained by two recent microbiological methods[16, 50] are not in good agreement. Some values are listed in TABLE 5. It would appear that values obtained with *Lactobacillus delbrueckii* are somewhat low, since they are below the amounts isolated, in some instances. *Leuconostoc mesenteroides*, on the other hand, gives values considerably higher than those reported by isolation, but in better agreement with isotope dilution data. Again, the value of comparing results obtained with two or more organisms is evident. Such comparisons have greater validity, however, if carried out in the same laboratory, since other factors which may affect the result are then constant. When tested in the same laboratory, these two test organisms gave identical results for the histidine content of two proteins, although these results were considerably higher than the "best values" given by Block and Bolling.[53] It is quite possible that the differences cited for aspartic acid reflect differences in sample treatment, method of hydrolysis, etc., in the two laboratories, rather than differences in the microbiological methods *per se*. With the availability of these new and rapid methods, the effect of variations in sample treatment on the assay values obtained should be determined, and techniques for preparation of samples standardized, in so far as possible.*

Since another author is to consider the amino acid composition of proteins, it does not seem fruitful, here, to compare further the results of traditional methods with those of microbiological assay methods. Such assay methods have been studied in some detail for leucine,[6, 19, 54] isoleucine,[19, 54] valine,[19, 39, 54] tryptophane,[4, 15, 55] serine,[16] glutamic acid,[56-59] aspartic acid,[16, 50] arginine,[39] and lysine.[45] Practically without exception, they appear to be capable of yielding precise results. Where the values obtained differ from older data, the microbiological methods, more often than not, have checked most closely the newer physico-chemical methods based on solubility product or isotope dilution measurements. This should make us hopeful for the future and more than ever interested in the refinement of these methods and their extension to additional amino acids.

In addition to the large number of amino acids which may be determined in this fashion, the methods have other advantages. The manual operations involved in setting up the various assays are very similar. Once the technique is learned for one amino acid, its application to others offers no additional problems. The equipment required is simple, comparatively inexpensive, readily available, and is the same for all amino acids. In most cases, especially with pure proteins, no elaborate or time-consuming procedures are necessary in preparing the sample for assay. One simply hydrolyzes the sample, neutralizes the excess acid, and dilutes for assay. It is, thus, possible, in a matter of weeks, to accumulate data which would require years of work by older methods. Finally, to one accustomed to working with older methods, the microbiological tests are almost unbelievably sensitive. The range over which the growth response of typical organisms increases with concentration of the amino acid, is shown in TABLE 6. Permitting four times as much of an amino acid, for the determination, as the maximum point on the assay range (an outside figure), and assuming a protein with a composition similar to that of casein, one may calculate that all eighteen amino acids could be determined on a 0.5 gm. sample of protein. Some investigators[39, 50, 59] have regularly used 2.5 cc. of medium, instead of 10 cc. On this scale, 125 mg. of protein would suffice for the analysis. Similar vitamin assays have been carried out on a scale 1/50 that commonly used (*i.e.,* 0.2 cc. total volume[60]), without sacrificing precision, accuracy, or speed. There is no apparent reason why this should not also be possible with the amino acid determinations. On this scale, the eighteen determinations could be made on a 10 mg. sample of protein. Less than half this amount would permit estimation of 13 of the 18. With development of these methods, the rarity of a protein or the cost of its preparation need no longer prevent its analysis.

---

* A valuable procedure for minimizing such differences between laboratories would be to include, with each assay for an individual amino acid, a sample in which content of that amino acid had been determined by a procedure of undoubted validity, such as the isotope dilution method. Failure to check the known value for this standard protein would immediately render valueless all other determinations. It has been suggested that such a standard protein be made available by some appropriate central agency. It is obvious, of course, that agreement with the known value for the standard protein would not guarantee accuracy of figures for other markedly different samples.

TABLE 6

ASSAY RANGES FOR THE VARIOUS AMINO ACIDS*

Amino acid	Assay range γ per 10 cc.	Amino acid	Assay range γ per 10 cc.
Aspartic acid	50–500	Leucine	10–100
Glutamic acid	50–500	Methionine	10–100
Arginine	10–80	Phenylalanine	10–150
Histidine	5–50	Proline	10–80
Lysine	10–200	Serine	50–250
Alanine	40–500	Threonine	50–500
Cystine	......	Tryptophane	2–25
Glycine	10–100	Tyrosine	5–80
Isoleucine	10–100	Valine	10–100

* The assay ranges given in the table are taken from publications referred to in the text. Wherever no detailed methods have been presented, the ranges are taken from the curves presented by Dunn *et al.*[25] showing the amino acid requirements of *Leuconostoc mesenteroides.* It may be expected that, as improved media for assay are developed, the ranges will be revised downwards rather than upwards.

So far, the specificity of response of these organisms to a required amino acid has not been considered. What effect do structures closely related to an amino acid have upon its determination? Almost all who have studied these assay methods have concerned themselves, to some extent, with this problem. In general, where complete media are used, the lactic acid bacteria have proved very specific in their response. With some exceptions, only the *l,* or naturally occurring, isomer possesses growth-promoting activity.[16, 61] Presence of the *d* form is without effect on the assay. Peptides are usually, but not always, less active than the equivalent quantity of the free amino acid. Complete hydrolysis of samples for assay is, therefore, generally necessary. Where structures other than the lacking amino acid are effective in promoting growth, it often indicates that the foreign structure can be converted to the amino acid by the test organism. Thus, in addition to providing necessary information for permitting specific estimation of an amino acid, studies of specificity of response sometimes contribute fundamental information concerning the formation or assimilation of an amino acid. Some examples may illustrate this point.

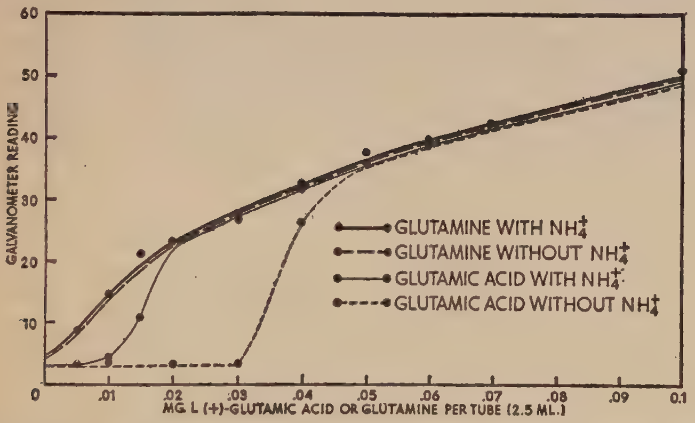
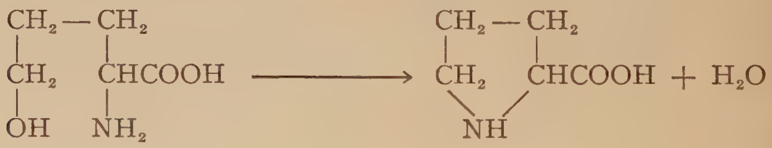


FIGURE 1. Comparative response of *L. arabinosus* to glutamine and *l*(+)-glutamic acid in the presence and absence of ammonium sulfate.

Anthranilic acid and indole were postulated as precursors of tryptophane, because they could be utilized in place of tryptophane by some bacteria.[62] This supposition was beautifully confirmed and extended by studies with *Neurospora* mutants, which demonstrated that biosynthesis of tryptophane took place *via* anthranilic acid and indole, the last stage consisting of the condensation of indole with serine, to form tryptophane.[63, 64]

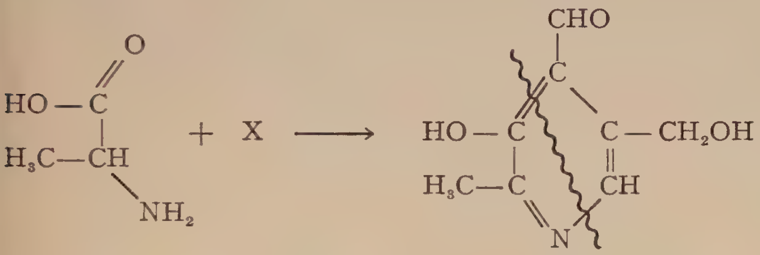
In the bioassay of glutamic acid, several investigators[57, 59] noted that the dose-response curve to glutamic acid was quite different from that obtained with other amino acids. Small additions of glutamic acid produced little or no increase in growth, while further additions produced a very rapid increase, followed by a more gradual rise. Glutamine was more active than glutamic acid, especially at low levels, and gave regularly ascending dose-response curves, of the type obtained with other amino acids. All modifications in the medium and conditions of growth which increased the availability of glutamic acid caused the dose-response curve more nearly to approach that of glutamine. It was concluded[58, 59] that at least a portion of the glutamic acid must be transformed to glutamine, for growth to occur. Data similar to those given in FIGURE 1 showed that ammonia aided in this conversion, because its presence greatly increased the availability of glutamic acid at low levels, but was without effect on the availability of glutamine.[59] Dunn and co-workers[56] reported no lag of this nature in the response of *Lactobacillus arabinosus* to increasing amounts of glutamic acid. Their medium contained asparagine in place of aspartic acid. Baumgarten and co-workers[59a] showed that addition of asparagine to the medium in place of aspartic acid prevented the sigmoidal curves otherwise obtained upon addition of glutamic acid. It is probable that the presence of asparagine promotes the formation of glutamine from glutamic acid, as do ammonium salts, and does so more effectively. Data of this type are of obvious value in permitting one to devise proper media and conditions for assay of an amino acid.

From the time of their original isolation, the atypical character of proline and hydroxyproline as amino acids gave rise to speculation as to whether these substances occurred, as such, in the protein molecule or whether they arose during hydrolysis from some labile precursor.[65] Sørenson later showed that α-amino-δ-hydroxyvaleric acid (but not the isomeric α-hydroxy-δ-aminovaleric acid) gave rise to proline on acid treatment; and pointed out that this product might occur in proteins



in place of, or together with, proline.[66] The natural occurrence of proline was later satisfactorily established by its isolation (in low yield) from enzymatic digests of protein.[67] The availability of an organism, *Leuconostoc mesenteroides,* for which proline is essential, permitted direct test of the physiological availability of α-amino-δ-hydroxyvaleric acid. Appropriate tests showed it to be completely inactive in promoting growth at levels 300 times those which showed some effect with proline (unpublished data).

One more example of the interesting findings which sometimes result from specificity studies will serve to conclude this discussion. *Streptococcus fecalis* R, an organism which ordinarily requires vitamin B_6 for growth, was found to grow in the absence of this vitamin, if sufficient alanine were present in the medium. Alanine was also an essential amino acid for growth in the presence of vitamin B_6.[22] It was postulated that alanine might serve as a precursor for synthesis of vitamin B_6 for this organism, in addition to being required for synthesis of protein. This suggested that an organism such as *Lactobacillus casei,* which could not use alanine in place of vitamin B_6 in media then available, might fail to do so for lack of a second precursor, as indicated below:



Alanine *Pyridoxal*

Although the truth or falsity of this hypothesis had not been established, a substance (X?) was found[68] in enzymatic protein digests, which permitted growth of *Lactobacillus casei* in the absence of vitamin B_6, if *dl*-alanine were also present in the medium. The unknown factor in enzymatic digests may be a hitherto-undiscovered amino acid or, more probably, a peptide. It is destroyed by both alkaline and acid hydrolysis. Next, the comparative activities of *l*- and *d*-alanine in permitting growth of these organisms in the presence and absence of vitamin B_6 was determined. The values found are given in TABLE 7. As expected, *l*(+)-alanine was more effective than its enantiomorph, in permitting growth of *Streptococcus fecalis* when vitamin B_6 was present in the medium; contrary to expecta-

TABLE 7

COMPARATIVE GROWTH-PROMOTING ACTIVITY OF *dl*-, *l*(+)-, AND *d*(—)-ALANINE FOR *S. fecalis* AND *L. casei* IN THE PRESENCE AND ABSENCE OF VITAMIN B_6[68]

Amino acid	Relative activity			
	Vitamin B_6 present		Vitamin B_6 absent	
	S. fecalis	*L. casei*	*S. fecalis*	*L. casei*
dl-Alanine	1.0	Not required	1.0	1.0
l(+)-Alanine	1.3		0.36	0.0075
d(—)-Alanine	0.33		2.0	2.0

tions, *d*(—)-alanine proved the more active in permitting growth in the absence of vitamin B_6. Indeed, for *Lactobacillus casei,* only *d*(—)-alanine has significant activity in this regard. This is more strikingly shown in FIGURE 2. This appears to be the first instance in which an amino acid of "unnatural" configuration has been shown to be essential for normal growth processes.* It would be of great interest to determine how it is acting, and whether other amino acids of the *d*-series may also play a role in metabolism.

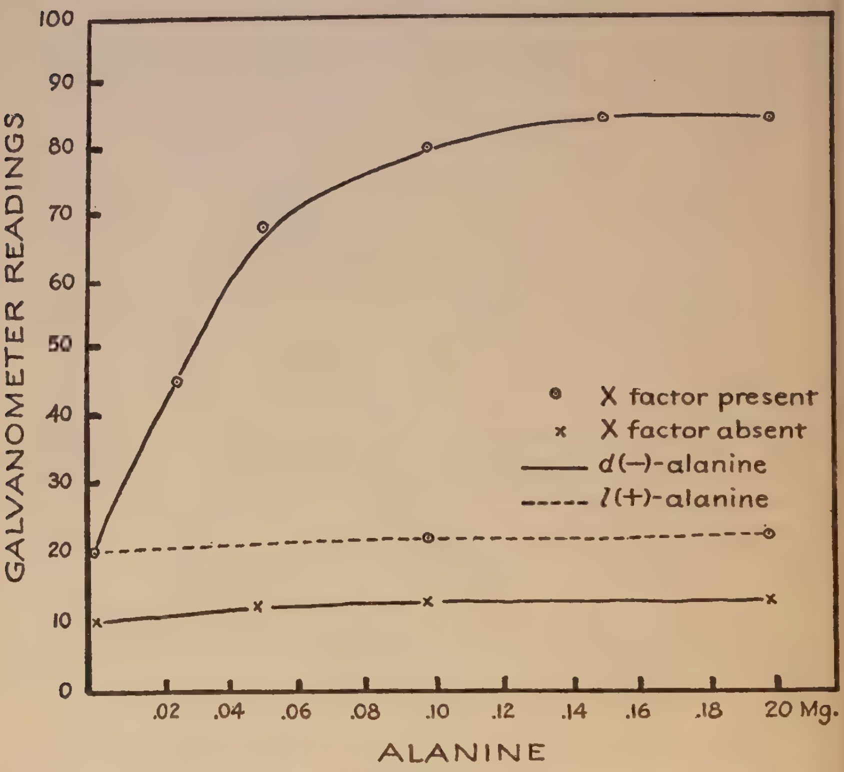


FIGURE 2. Response of *Lactobacillus casei* to *d*(—)- and *l*(+)-alanine in the presence and absence of an enzymatic casein digest (X — factor).

From the results cited above, and others in the literature, it is obvious that the lactic acid bacteria can be used successfully for the accurate and convenient estimation of many amino acids. Their use for this purpose is being rapidly extended. The author believes that these methods are destined to play a large, if not a predominant role in the determination of amino acids, for many years to come. In addition, they can furnish interesting and perhaps useful information concerning the biogenesis, mode of assimilation, and function of various amino acids. The methods had their origin in the detailed study of bacterial nutrition. This study is of recent origin, and further investigation along this line should permit considerable improvement of methods now available, and the development of many additional ones. Such study clearly merits the interest and encouragement of all protein chemists.

---

* Dunn *et al.*[50] had previously postulated that *d*(—)-glutamic acid, though not essential for growth, played some role in the normal metabolism of *L. arabinosus,* since *dl*-glutamic acid appeared to permit higher acid production, when present in excess, than did *l*(+)-glutamic acid. Hac *et al.*,[59] using turbidity measurements, were unable to confirm this suggestion.

BIBLIOGRAPHY

1. **Snell, E. E.**
   1945. Advances in Protein Chemistry **II**: 85–118. Academic Press. New York.
2. **Snell, E. E., & F. M. Strong**
   1939. Ind. Eng. Chem. Anal. Ed. **11**: 346.
3. **Snell, E. E., & L. D. Wright**
   1941. J. Biol. Chem. **139**: 675.
4. **Wooley, J. G., & W. H. Sebrell**
   1945. J. Biol. Chem. **157**: 141.
5. **Beadle, G. W., & E. L. Tatum**
   1941. Proc. Nat. Acad. Sci. **27**: 499.
6. **Ryan, F. J., & E. Brand**
   1944. J. Biol. Chem. **154**: 161.
7. **Roepke, R. R., R. L. Libby, & M. H. Small**
   1944. J. Bact. **48**: 401.
8. **Hunter, A., & J. A. Dauphinee**
   1930. J. Biol. Chem. **85**: 627.
9. **Zittle, C. A., & N. R. Eldred**
   1944. J. Biol. Chem. **156**: 401.
10. **Baddiley, J., & E. F. Gale**
   1945. Nature **155**: 727.
11. **Umbreit, W. W., & I. C. Gunsalus**
   1945. J. Biol. Chem. **159**: 333.
12. **Snell, E. E., B. M. Guirard, & R. J. Williams**
   1942. J. Biol. Chem. **143**: 519.
13. **Niven, C. F., Jr.**
   1944. J. Bact. **47**: 343.
14. **Snell, E. E., & A. N. Rannefeld**
   1945. J. Biol. Chem. **157**: 475.
15. **Greene, R. D., & A. Black**
   1944. J. Biol. Chem. **155**: 1.
16. **Stokes, J. L., & M. Gunness**
   1945. J. Biol. Chem. **157**: 651.
17. **Shankman, S.**
   1943. J. Biol. Chem. **150**: 305.

18. **Hegsted, D. M.**
1944. J. Biol. Chem. **152**: 193.

19. **Kuiken, K. A., W. H. Norman, C. M. Lyman, F. Hale, & L. Blotter**
1943. J. Biol. Chem. **151**: 615.

20. **Hutchings, B. L., & W. H. Peterson**
1943. Proc. Soc. Exp. Biol. & Med. **52**: 36.

21. **Stokes, J. L., & M. Gunness**
1945. Science **101**: 43.

22. **Snell, E. E., & B. M. Guirard**
1943. Proc. Nat. Acad. Sci. **29**: 66.

23. **Hegsted, D. M., & E. D. Wardwell**
1944. J. Biol. Chem. **153**: 167.

24. **Niven, C. F., Jr.**
1944. J. Bact. **47**: 343.

25. **Dunn, M. S., S. Shankman, M. N. Camien, W. Frankl, & L. B. Rockland**
1944. J. Biol. Chem. **156**: 703.

26. **Kidder, G. W., & V. C. Dewey**
1945. Arch. Biochem. **6**: 425.

27. **Wright, L. D., & H. R. Skeggs**
1945. J. Biol. Chem. **159**: 611.

28. **Snell, E. E.**
1944. J. Biol. Chem. **154**: 313.

29. **Gladstone, G. P.**
1939. Brit. J. Exp. Path. **20**: 189.

30. **Pelczar, M. J., & J. R. Porter**
1943. Arch. Biochem. **2**: 323.

31. **Woolley, D. W.**
1941. J. Biol. Chem. **140**: 453.

32. **Dunn, M. S., S. Shankman, M. N. Camien, & L. B. Rockland**
1944. J. Biol. Chem. **156**: 703.

33. **Krehl, W. A., F. M. Strong, & C. A. Elvehjem**
1943. Ind. Eng. Chem. Anal. Ed. **15**: 471.

34. **Lampen, O.**
1945. Report to AAAS Conference on Food and Nutrition, Gibson Island, Md.

35. **Atkin, L., A. S. Schultz, W. L. Williams, & C. N. Frey**
1943. Ind. Eng. Chem. Anal. Ed. **15**: 141.

36. **Stokes, J. L., A. Larsen, C. R. Woodward, Jr., & J. W. Foster**
1943. J. Biol. Chem. **150**: 17.

37. **Wood, E. C.**
1945. Nature **155**: 632.

38. **Leonian, L. H., & V. G. Lilly**
1945. Agric. Exp. Station, West Virginia Univ. Bull. 319.

39. **McMahan, J. R., & E. E. Snell**
1944. J. Biol. Chem. **152**: 83.

39a. **Guirard, B. M., E. E. Snell, & R. J. Williams**
1946. Proc. Soc. Exp. Biol. & Med. **61**: 158.

40. **Vickery, H. B.**
1940. J. Biol. Chem. **132**: 325.

41. **Foster, G. L.**
1945. J. Biol. Chem. **159**: 431.

42. **Chibnall, A. C.**
1942. Proc. Roy. Soc. London B **131**: 136.

43. **Bergman, M., & C. Niemann**
1938. J. Biol. Chem. **122**: 577.

44. **Neuberger, A.**
1945. J. Biol. Chem. **158**: 717.

45. **Dunn, M. S., M. N. Camien, S. Shankman, W. Frankl, & L. B. Rockland**
1944. J. Biol. Chem. **156**: 715.

46. **Vickery, H. B., & A. White**
1933. J. Biol. Chem. **103**: 413.

47. **Cannan, R. K., A. H. Palmer, & A. C. Kibrick**
1942. J. Biol. Chem. **142**: 803.

48. **Aberhalden, E.**
1922. Z. physiol. Chem. **142**: 803.

49. **Vickery, H. B., & R. J. Block**
1931. J. Biol. Chem. **93**: 105.

50. **Hac, L. R., & E. E. Snell**
1945. J. Biol. Chem. **159**: 291.

51. **Bailey, K., A. C. Chibnall, M. W. Rees, & E. F. Williams**
1943. Biochem. J. **37**: 360.

52. **Chibnall, A. C., M. W. Rees, & E. F. Williams**
1943. Biochem. J. **37**: 371.

53. **Block, R. J., & D. Bolling**
1945. The Determination of the Amino Acids. Springfield, Ill.

54. **Schweigert, B. S., J. M. McIntire, C. A. Elvehjem, & F. M. Strong**
1944. J. Biol. Chem. **155**: 183.

55. **Dunn, M. S., H. F. Schott, W. Frankl, & L. B. Rockland**
1945. J. Biol. Chem. **157**: 387.

56. **Dunn, M. S., M. N. Camien, L. B. Rockland, S. Shankman, & S. C. Goldberg**
1944. J. Biol. Chem. **155**: 591.

57. **Lewis, J. C., & H. S. Olcott**
1945. J. Biol. Chem. **157**: 265.

58. **Lyman, C. A., K. A. Kuiken, L. Blotter, & F. Hale**
1945. J. Biol. Chem. **157**: 395.

59. **Hac, L. R., E. E. Snell, & R. J. Williams**
1945. J. Biol. Chem. **159**: 273.

59a. **Baumgarten, W., A. N. Mather, & W. Stone**
1945. Cereal Chem. **22**: 514.

60. **Lowry, O. H., & O. A. Bessey**
1944. J. Biol. Chem. **155**: 71.

61. **Stokes, J. L., & M. Gunness**
1944. J. Biol. Chem. **154**: 715.

62. **Snell, E. E.**
1943. Arch. Biochem. **2**: 389.

63. **Tatum, E. L., D. M. Bonner, & G. W. Beadle**
1944. Arch. Biochem. **3**: 477.

64. **Tatum, E. L., & D. M. Bonner**
1944. Proc. Nat. Acad. Sci. **30**: 30.

65. **Fischer, E.**
1901. Z. physiol. Chem. **33**: 151.

66. **Sørenson, S. P. L.**
1905. Compt. rend. Trav. Lab. Carlsberg **6**: 137.

67. **Fischer, E., & E. Abderhalden**
1904. Z. physiol. Chem. **40**: 215.

68. **Snell, E. E.**
1945. J. Biol. Chem. **158**: 497.

ENZYMATIC METHODS IN AMINO ACID ANALYSIS

By Reginald M. Archibald

Hospital of The Rockefeller Institute for Medical Research, New York, N.Y.

Except for the so-called microbiological methods already considered in these papers, relatively little use has been made of specific enzymes for the determination of amino acids in protein hydrolysates.

The enzymatic methods which, so far, have been worked out in detail are: (a) The arginase method for arginine; (b) the decarboxylase method for lysine, ornithine, tyrosine, arginine, glutamic acid, and histidine; and (c) Cohen's method[1] for glutamic acid, which involves conversion of glutamic acid to succinic acid, by the action of chloramine T, followed by an enzymatic measurement of the succinic acid by the use of its specific dehydrogenase.

The first two of these methods deserve further comment.

Arginase Methods

The arginase method, as worked out by Hunter and his colleagues,[2, 3] is fairly simple. It is reliable and accurate. Because some of the amino acids, especially lysine, present in protein hydrolysates strongly inhibit the action of arginase,[4] the incubation period of the enzyme is necessarily long. By the use of manganese-activated preparations of arginase,[5, 6] the necessary incubation period can be cut down from 12 hours to 2 hours.

Arginine is hydrolyzed by arginase to yield ornithine and urea. It is conveniently and reliably measured by any one of three methods: (a) colorimetric measurement of the urea produced (0.02 to 0.2 mg. sample of arginine);[5, 7] (b) manometric measurement of the CO_2 liberated by the action of urease on the urea formed (8 mg. sample);[3] (c) titrimetric[2] or colorimetric measurement of the NH_3 liberated by the action of urease on the urea (8 mg. sample). The last two methods have been simplified by the use of dialyzed urease which is free of canavanine.[5, 8]

The arginase methods are specific for arginine. Interference is possible from canavanine, which also is hydrolyzed to yield urea, though much more slowly than is arginine. So far, however, the presence of canavanine in protein hydrolysates has not been demonstrated. In summary, of the arginase methods it may be said that they are accurate, specific, reliable, relatively simple, and moderately rapid.

Decarboxylase Methods

Recently, increased interest in the biochemistry of microorganisms has led to the discovery of some very promising analytical tools, *viz.*, the amino acid decarboxylases. Numerous microorganisms when grown in acidic media have been found to elaborate one or more decarboxylases, each of which is specific for a given amino acid. In the presence of either a natural[9] or readily synthesized[10, 11] coenzyme (pyridoxal phosphate),* any one of these decarboxylases liberates CO_2 from the carboxyl group of the amino acid for which it is specific. Gale and Epps[12] have prepared from *B. cadaveris* a decarboxylase which is specific for *l* (+)-lysine, except for a very slow decomposition of hydroxylysine.[13] An arginine decarboxylase present in the organism is destroyed during the purification process. These workers observed that *Escherichia coli* contained decarboxylases for arginine, lysine, histidine, and ornithine. The ornithine enzyme in this organism could be destroyed by acetone drying, but the histidine and arginine enzymes could be removed from the lysine enzyme only in part. Zittle and Eldred[14] dispensed with elaborate purification procedures and employed month-old preparations of acetone dried organisms of *Bacterium cadaveris*. They found these to be free of arginine decarboxylase and used them for the determination of lysine in protein hydrolysates. Epps[15] has prepared from *Streptococcus fecalis* a decarboxylase which is specific for *l*(—)-tyrosine and *l*-3,4-dihydroxyphenylalanine. Separate decarboxylases for these two amino acids are to be found also in animal tissue.[16]

Epps[17] prepared a specific *l*(—)-histidine decarboxylase from *Clostridium welchii.* The *l*(+)-glutamic acid decarboxylase, which was also present in the organism, was destroyed during purification of the histidine decarboxylase.

Gale observed[18] that one strain of *Clostridium welchii* contained an enzyme which would decarboxylate only *l*(+)-glutamic acid and β-hydroxy-glutamic acid. He employed thick suspensions of the intact organism for the determination of glutamic acid in protein hydrolysates.

Gale observed, further, that the only amino acid decarboxylated by a washed suspension of *Clostridium septicum* was *l*(+)-ornithine. He[18] employed preparations of the specific decarboxylases of lysine, histidine, tyrosine, glutamic acid, and ornithine, for the determination of these amino acids in acid hydrolysates of 8 to 10 proteins.

From the literature, it would appear that, thus far, all workers have conducted their decarboxylations in a Warburg manometer. It is probable that the procedure can be adapted readily to the use of the Van Slyke–Neill blood gas apparatus.[19] This would facilitate the determinations for a number of workers not supplied with Warburg manometers. Neuberger and Sanger[20] observed that 1–2 mg. of lysine in a casein hydrolysate could be decarboxylated completely, in 5 minutes, at 37° C. in a Warburg apparatus.

---

* It appears likely that histidine decarboxylase and glutamic acid decarboxylases do not require the codecarboxylase which activates the decarboxylases of arginine, lysine, tyrosine and ornithine.[22]

Brand[21] has observed that Gale's values for lysine, by the decarboxylase method, tend to be low. Brand, therefore, prefers his bioassay for the determination of lysine. Values obtained by the decarboxylase method may become theoretical, when smaller gas solubility correction factors are required, as would be the case if a Van Slyke gas apparatus were employed. Preliminary work[5] with decarboxylases in the Van Slyke machine indicates that these enzymes measure accurately and specifically the *l*-form of glutamic acid, lysine, or arginine in protein hydrolysates, and in blood plasmas. One danger inherent in the use of the decarboxylase methods is that the cultures, from which the enzyme preparations are obtained, may in time change their specificity. An example of such a change in a strain of *Escherichia coli* already appears to have been recorded in the literature.[11] It will be necessary, therefore, to check the specificity of new enzyme or bacterial preparations, from time to time.

Specificity may be achieved in some preparations containing two or more decarboxylases, by the use of selective inhibitors such as hydrazine or semicarbazide.[22]

In summary, the decarboxylase methods show great promise. They appear to be specific and, once the enzyme preparations are on hand, simple to use. They are rapid and appear capable of accuracy.

So much for the enzymatic methods which have proven their worth when applied to protein hydrolysates. Now I shall discuss enzymatic methods which, so far, have *not* been worked out in detail for application to hydrolysates, but which should be borne in mind for future development. They may be able to provide important and accurate data, which would be difficult to obtain by the non-enzymatic methods now available.

Amidases

An accurate enzymatic method[23, 24] is available for the determination of glutamine in mixtures of other amino acids. Geddes and Hunter[25] have obtained from yeast an enzyme preparation which contains asparaginase and glutaminase. Enzymatic determinations of glutamine and asparagine are, therefore, possible. Asparagine would be determined by subtracting the glutamine value from that obtained for the sum of the two amides. However, the acid hydrolysis of proteins is accompanied by a hydrolysis of the amide group of the dicarboxylic acid amides. Hence, there is neither glutamine nor asparagine in the acid hydrolysates. There is, therefore, no point in applying the glutaminase or the asparaginase method to such hydrolysates.

If, however, there can be found a combination of proteolytic enzymes which is capable of hydrolyzing all the peptide linkages of protein and which, at the same time, is free of amidases, then the methods for the determination of glutamine and asparagine would be applicable to the digest. At present, no such combination of enzymes is known. It is possible, however, that, even if active proteolytic preparations are not freed entirely of amidases, the enzymatic hydrolysis of amides may be prevented, during peptide fission, by the addition of a compound such as bromsulfalein, which has been shown to be a strong inhibitor of glutaminase.[23] In this case, the enzymatic digest subsequently could be subjected to analysis for glutamine by non-enzymatic methods, such as the ninhydrin method of Hamilton.[26] Alternatively, the inhibitor might be removed by extraction, following which the enzymatic amidase methods could be employed.

Unless the enzymatic fission of peptide linkages can be accomplished rapidly and at pH's, let us say, between 5 and 8, a considerable portion of the rather labile glutamine will have been hydrolyzed before the digestion is complete. It appears likely, therefore, that the distribution of amide nitrogen between the glutamine and asparagine of proteins will be ascertained, either by an analysis of the rate curves of the hydrolysis of amide N (anionic detergents might be employed as catalysts, in the manner described by Steinhardt and Fugitt),[27] or by application of non-enzymatic methods of measuring the two amides in enzymatic hydrolysates of proteins. The determination of the distribution of amide nitrogen between the two dicarboxylic amino acids in proteins remains, then, one of the difficult and unsolved problems of protein chemistry.

d-Amino Acid Oxidase

The use of *d*-amino acid oxidase for the determination of *d*-amino acids has seldom been applied to the quantitative analysis of a protein hydrolysate. The method, however, should not be overlooked entirely. Preparations of the *d*-oxidase could be freed of *l*-oxidase activity, either by drying or by inhibition of the *l*-oxidase with cyanide or octyl alcohol.[28] Such a preparation, applied to a protein hydrolysate, could give a value for the total of those *d*-forms of amino acids present which were susceptible to oxidation by the *d*-amino acid oxidase. If this analysis failed to detect any of the *d*-form, considerable time could then be saved, by avoiding the determination (by the isotope dilution methods) of such of the *d*-forms as are known to be deaminated, at a rapid or moderately rapid rate,[29] by the *d*-oxidase.

BIBLIOGRAPHY

1. **Cohen, P. P.**
   1939. Biochem. J. **33**: 551.
2. **Hunter, A., & J. A. Dauphinee**
   1930. J. Biol. Chem. **85**: 627.
3. **Hunter, A., & J. B. Pettigrew**
   1937. Enzymologia **1** (6): 341.
4. **Hunter, A., & C. E. Downs**
   1945. J. Biol. Chem. **157**: 427.
5. **Archibald, R. M.**
   Unpublished.
6. **Van Slyke, D. D., & R. M. Archibald**
   1942. Fed. Proc. **1**: 139.
7. **Archibald, R. M.**
   1945. J. Biol. Chem. **157**: 507.
8. **Archibald, R. M., & P. B. Hamilton**
   1943. J. Biol. Chem. **150**: 155.
9. **Gale, E. F., & H. M. R. Epps**
   1944. Biochem. J. **38**: 250.
10. **Baddiley, J., & E. F. Gale**
    1945. Nature **155**: 727.
11. **Umbreit, W. W., & I. C. Gunsalus**
    1945. J. Biol. Chem. **159**: 333.
12. **Gale, E. F., & H. M. R. Epps**
    1944. Biochem. J. **38**: 232.
13. **Gale, E. F., & H. M. R. Epps**
    1943. Nature **152**: 327.
14. **Zittle, C. A., & N. R. Eldred**
    1944. J. Biol. Chem. **156**: 401.
15. **Epps, H. M. R.**
    1944. Biochem J. **38**: 242.
16. **Blaschko, H.**
    1942. J. Physiol. **101**: 337.
17. **Epps, H. M. R.**
    1945. Biochem. J. **39**: 42.
18. **Gale, E. F.**
    1945. Biochem. J. **39**: 46.
19. **Van Slyke, D. D.**
    1927. J. Biol. Chem. **73**: 121.
20. **Neuberger, A., & F. Sanger**
    1944. Biochem. J. **38**: 119.
21. **Brand, E.**
    1945. J. Am. Chem. Soc. **67**: 1532.

22. **Taylor, E. S., & E. F. Gale**
1945. Biochem. J. **39**: 52.
23. **Archibald, R. M.**
1944. J. Biol. Chem. **154**: 657.
24. **Archibald, R. M.**
1944. J. Biol. Chem. **154**: 643.
25. **Geddes, W. F., & A. Hunter**
1928. J. Biol. Chem. **77**: 197.
26. **Hamilton, P. B.**
1945. J. Biol. Chem. **158**: 375.
27. **Steinhardt, J., & C. H. Fugitt**
1942. J. Res. Nat. Bur. Standards **29**: 315.
28. **Krebs, H. A.**
1935. Biochem J. **29**: 1620.
29. **Horowitz, N. H.**
1944. J. Biol. Chem. **154**: 141.

AMINO ACID COMPOSITION OF SIMPLE PROTEINS

By Erwin Brand

College of Physicians and Surgeons, Columbia University, New York, N.Y.

From the point of view of organic chemistry, the establishment of the empirical formula (cf. Mulder,[1] in 1838) of proteins, in terms of constituent amino acids, is fundamental to an understanding of their constitution and to the development of theories of protein structure.

An apparently complete analysis of a simple, crystalline, homogeneous protein of small molecular weight (42,000), β-lactoglobulin, has recently been presented.[2] Although, doubtless, there will be future improvements in detail, this paper is the first to describe a complete analysis of a single preparation of a protein by the same group of workers.

Many of the data were obtained in the course of the development and standardization of methods for application to plasma proteins. This represents part of the comprehensive program of the Plasma Fractionation Laboratory, at Harvard University, under the direction of Dr. Edwin J. Cohn. A considerable part of the work presented here was carried out under contract, recommended by the Committee on Medical Research, between the Office of Scientific Research and Development and Columbia University.

Data on β-lactoglobulin, bovine serum albumin, human γ-globulin, and crystalline insulin will be given. Alanine is the only amino acid which has not been determined in these proteins (except in β-lactoglobulin). All the values given in these tables are our own results.*

METHODS OF ANALYSIS

Our methods are summarized in TABLE 1. Some of these methods have been reported in detail; others, in outline; and some are, as yet, unpublished. They cover a wide range of chemical manipulations, from ultraviolet absorption, micro-diffusion, micro-titration, and photometry to micro bioassays, with *Neurospora* mutants and different strains of *Lactobaccilli,* and enzymatic methods with the specific decarboxylase of *Bacterium cadaveris*. The development of the bioassays was possible as the result of the recent pioneer work of Beadle, Snell, Dunn, Shankman, Lyman, Lewis, Olcott, Gale (decarboxylases), and others.

---

* In the development of the work, Dr. Beatrice Kassell, Dr. Leo J. Saidel, Mr. W. H. Goldwater, Mr. E. M. Diskant, and Dr. Francis J. Ryan, from the Department of Zoology, have taken part for the past few years and much credit is due to them.

TABLE 1

METHODS OF ANALYSIS

Det. No.	Constituent	Method	Authors	Ref. No.
1	Carbon	Pregl.	Saschek & Brand	54
2	Hydrogen	“	“ “ “	54
3	Total Nitrogen	Micro Dumas, preferred		
4	Total Sulfur	Pregl.	Saschek Kassell & Brand	53 45
5	Oxygen	Difference: Sum of Oxygen of AA residues	Brand *et al.*	3
7	Sulfate S	Gravimetric, Iodometric	Kassell & Brand	45
8	Amide N	Micro-diffusion	Saidel & Brand	55 (cf. 56)
11	Total Free Amino N	Van Slyke, man.	Diskant & Brand	57
12	Free α-Amino N	11 less ε-Amino N	Brand *et al.*	3
14	Glycine	Bioassay, LM	Brand *et al.*	7
15	Alanine	“ SF	“ “ “	7
16	Valine	“ LA, LM	“ “ “	7
17	Leucine	“ NL	Ryan & Brand	12
	“	“ LA, LM	Brand *et al.*	7
18	Isoleucine	“ LA, LM	“ “ “	7
19	Proline	“ NP, LM	Ryan & Brand	3, 7
20	Phenylalanine	“ LA, LM	Brand *et al.*	7
21	Cysteine	Photometric, Iod.	Kassell & Brand	45, 50, 52
22	Half-Cystine	“ “	“ “ “	45, 50, 52
23	Methionine	Iodometric	“ “ “	45
24	Tryptophane	Ultraviolet Abs.	Brand & Saidel	47 (cf. 46)
25	Arginine	Photometric	Brand & Kassell	24
26	Histidine	“	Saidel & Brand	36
27	Lysine	Bioassay, LM, BC	Brand *et al.*	3, 7, 26
28	Aspartic Acid	“ LM	“ “ “	7
29	Glutamic Acid	“ LA	“ “ “	7
32	Serine	Corrected for 10% destruction	“ “ “	3, 7
		Bioassay, LM	“ “ “	7
		Oxidation	“ “ “	3, 7
			cf. Nicolet *et al.*	49
33	Threonine	Oxidation	Brand *et al.*	3, 7
			cf. Nicolet	49, 58
			cf. Winnick	59
34	Tyrosine	Photometric	Brand & Kassell	46
35	Hydroxyproline	Colorimetric	McFarlane & Guest	60
37	“Terminal H_2O”	Corresponding to No. 12 N of AA	Brand *et al.*	3
38	Total AA Sum		Brand *et al.*	3
39	Water taken up on hydrolysis	By difference	“ “ “	3

TABLE 1 (Continued)

METHODS OF ANALYSIS

Moisture Determinations*

The proteins are equilibrated and the moisture determined *in vacuo* at 100° C. All determinations are carried out on the moisture-containing proteins and are subsequently corrected to dry weight.

Hydrolysis

1) 6 N HCl
2) 6 N H_2SO_4 for histidine[36]
3) HI for cysteine, methionine, and sulfate[45, 50, 52]
4) HCl-Urea for photometric cysteine and cystine[50, 52]
5) NaOH-$SnCl_2$ for tyrosine and tryptophane[46, 47]

Organisms used in Bioassays

NL = "Leucine-less" *Neurospora* mutant
NP = "Proline-less" *Neurospora* mutant
LA = *Lactobacillus arabinosus* 17-5
LC = *Lactobacillus casei* Σ
LD = *Lactobacillus delbrueckii*
LM = *Leuconostoc mesenteroides* P-60
SF = *Streptococcus fecalis* R
BC = Lysine decarboxylase of *Bacterium cadaveris*

* Cf. Brand & Kassell.[52]

In the micro bioassays (cf. TABLE 1), we have used both lactic acid titrations and turbidity measurements (with the Beckman Spectrophotometer), but have come to prefer the latter. The time for the micro bioassay with *Leuconostoc mesenteroides* has been shortened to about 22 hours (cf. TABLE 2). With our assay medium,[3] growth of this organism is practically complete after about 20 hours.[3]

In TABLE 2, part of a detailed protocol for a bioassay with *Leuconostoc mesenteroides* for lysine, in a β-lactoglobulin hydrolysate, is presented. The agreement of duplicate or triplicate tubes is usually very good.

A comparison of the results obtained by the newer methods of amino acid analysis is given in TABLE 4 for a number of proteins. The methods considered are briefly summarized in TABLE 3; + or —, in this table, indicates the presence or absence of values, by the methods tabulated for any one of the amino acids.

Values by the isotope dilution method are now available for glycine, leucine, phenylalanine (only one value reported), glutamic acid, aspartic acid, arginine, lysine, and tyrosine. Results by the microbiological methods are included for all the amino acids (for alanine, only one value), except for tyrosine, tryptophane, and cystine. In some cases, values obtained

Table 2

Bioassay for Lysine

Part of Protocol of Experiment No. 15 5/30/45
Medium No. 64. *Leuconostoc mesenteroides* P-60*

3 cc. of medium + 3 cc. of standard or of hydrolysate + 1 drop of inoculum were grown in 15 cc. centrifuge tubes at 36.8 ± 0.5 C. for 22 hours. The turbidity of the bacterial suspension was measured in a Beckman spectrophotometer at 6500 A in 1 cm. cells.

l-Lysine Standards		*β-Lactoglobulin Hydrolysate*		
γ/6 cc.	Extinction Coefficient	γ Protein per 6 cc.	Extinction Coefficient	Lysine Yield
	K × 10^3		K × 10^3	per cent
0	20			
	16			
		349	193	11.9
19.9	114		196	
	115			
		526	269	11.6
39.9	190		268	
	190		265	
59.9	274	701	328	11.2
	265		327	
	268		328	
79.8	340	875	396	11.2
	330		390	
	328		392	
99.6	400	1050	476	11.5
	395		476	
	401		475	
119.8	478	1226	538	11.0
	471		528	
	469		538	
140.5	567	Average		11.4 ± 0.1
	570	Isotope Dilution Method†		11.3
160.3	611			
	620			

* Culture obtained from Dr. E. E. Snell.
† Foster.[8]

TABLE 3

KEY TO AMINO ACID METHODS
(DATA IN TABLE 4)

For each method, + indicates a number of values for different proteins in TABLE 4; (+) indicates one value only; and —, the absence of data. Several + signs indicate modifications of the same type of method (*e.g.*, values obtained with different organisms in bioassays).

	Isotope Dilution	Solubility Product	Bioassay	Chromatography
Glycine	+	+	+	—
Alanine	—	—	(+)	+
Valine	—	—	++	+
Leucine	+	+	++++	(+)
Isoleucine	—	—	+++	(+)
Leucine and isoleucine	—	—	++++	+
Phenylalanine	(+)	—	+++	+
Proline	—	(+)	+(+)	+

	Isotope Dilution	Isolation	Bioassay	Photometry	Decarboxylase
Glutamic Acid	+	+	+	—	—
Aspartic Acid	+	+	+	—	—
Arginine	(+)	++	++	+++	+
Histidine	—	+	+++	+++	+
Lysine	+	++	++	—	+
Tyrosine	+	—	—	+	—
Tryptophane	—	—	—	++	—

	Oxidation	Iodometry	Bioassay	Photometry
Serine	++	—	+	—
Threonine	+++	—	+	—
Methionine	—	++	+	—
Cystine	—	+	—	+

with 3 or 4 different organisms are given in TABLE 4 (indicated by + signs in TABLE 3).

For alanine, valine, the leucines, phenylalanine, and proline, results obtained by partition chromatography are included in TABLE 4. This application of the chromatographic method to protein hydrolysates was recently introduced by Martin and Synge.[4] The data quoted, obtained by this

TABLE 4

COMPARISONS OF VALUES FOR THE AMINO ACID CONTENT OF PROTEINS

References are to the methods, as listed in the Bibliography for tables, and TABLE 1. The figures quoted are gm. of amino acid per 100 gm. of protein.

Explanation of the abbreviations for micro-organisms used in bioassays is given in TABLE 1.

The values originally reported by A. C. Chibnall and his co-workers* as per cent amino acid N of protein-N have been recalculated on a basis of gm. of amino acid per 100 gm. of protein. Total N of the proteins concerned is:† β-lactoglobulin 15.58 per cent, insulin C 15.54 per cent, egg albumin 15.76 per cent, edestin 18.65 per cent, horse hemoglobin 16.8 per cent (not corrected for ash), gelatin 18.0 per cent, casein 15.65 per cent.

Protein	Partition Chromatography	Microbiological Assay	Solubility Product	Isotope Dilution
Glycine				
β-Lactoglobulin		1.4 LM[3]		1.5[8]
Insulin C			4.3[18]	
Insulin V		4.6 LM[7]		
Horse Hemoglobin				5.6[8]
Human Serum Albumin		1.6 LM[7]		1.55[21]
Bovine Serum Albumin		1.9 LM[7]		1.96[21]
Bovine Fibrin		5.8 LM[7]		5.70[19]
Egg Albumin		3.0 LM[7]	3.21[13, 15]	
Gelatin		27.3 LM[7]	26.5[15, 16]	
Casein		2.7 LM[7]		
Alanine				
β-Lactoglobulin	6.64[1]	6.2 SF[3]		
Insulin C	4.60[1]			
Insulin V‡				
Egg Albumin	6.72[1]			
Edestin	4.31[1]			
Gelatin	9.27[1]			
Human Fibrin	3.7[4]			
Casein	3.19[1]			

* Cf.[1, 3, 4, 9, 18, 25, 26, 29, 30, 48]

† Cf.[2]

‡ 2.9 per cent, calculated by difference: Cf. TABLE 10 D.

TABLE 4 (Continued)

Protein	Partition Chromatography	Microbiological Assay	Solubility Product	Isotope Dilution
		Valine		
β-Lactoglobulin	5.34[1]	5.83 LA,[8] 5.5 SF,[10] 5.8 LC[5]		
Insulin C	7.49[1]			
Insulin V		8.8 LA[7]		
Horse Hemoglobin	8.42[1]	8.8 LC,[5] 10.7 LA[7]		
Egg Albumin	4.62[1]	6.8 LC,[5] 7.0 SF[10]		
Edestin	4.75[1]			
Gelatin	3.76[1]	2.7 LC,[5] 2.7 SF[10]		
Casein	6.02[1]	6.7 LC,[5] 6.86 LA[5] 6.25 LA,[6] 6.7 SF,[10] 7.4 LA[7]		
Lactalbumin		6.0 LA[7]		
Human Fibrin	3.9[4]	4.7 LA[7]		
		Leucine		
β-Lactoglobulin		15.4 NL,[12] 15.3 SF[10]	15.9[22]	15.7[8]
Insulin V		13.4 NL[12]		
Horse Hemoglobin		15.7 NL[12]		15.1[8]
Egg Albumin		9.6 NL,[12] 9.2 SF[10]	9.1[18]	
Edestin		7.4 NL[11]		
Gelatin		3.6 NL,[12] 3.30 LA[6] 3.5 SF[10]	3.5[13]	
Human Fibrin		7.1 NL[11]		
Casein	10.55[9]	9.8 NL,[12] 9.27 LA[6] 9.9 SF,[10] 9.6 LA[13]		
Lactalbumin		9.9 LM[7]		
		Isoleucine		
β-Lactoglobulin		7.0 SF,[10] 8.4 LA[8]		
Insulin V		2.9 LA[7]		
Horse Hemoglobin		0 LA[7]		
Egg Albumin		7.0 SF[10]		
Edestin				
Gelatin		1.72 SF,[10] 1.71 LA[6] 1.81 LA[7]		
Human Fibrin		5.2 LA[7]		
Casein	5.27[9]	5.6 SF,[10] 6.05 LA[6] 6.2 LA[7]		
Lactalbumin		6.2 LM[7]		
		Leucine + Isoleucine		
β-Lactoglobulin	20.1[1]	22.3 SF[10] 23.8 LA, NL[3, 12]		
Insulin V		16.3 LA, NL[7, 12]		
Insulin C	15.68[1]			
Egg Albumin	15.94[1]	16.2 SF[10]		
Edestin	12.13			
Gelatin	5.23[1]	5.41 LA, NL[7, 12] 5.01 LA,[6] 5.22 SF[10]		
Human Fibrin	11.7[4]	12.3 LA, NL[7, 11]		
Casein	15.82[1]	16.0 LA, NL[7, 12] 15.5 SF,[10] 15.32 LA[6]		
Lactalbumin		16.1 LM[7]		

TABLE 4 (Continued)

Protein	Partition Chromatography	Microbiological Assay	Solubility Product	Isotope Dilution
		Phenylalanine		
β-Lactoglobulin	4.60[1]	3.54 LA,[3] 4.3 LD[10]		
Insulin C	8.09[1]			
Horse Hemoglobin	6.53[1]	7.5 LA[7]		6.8[8]
Egg Albumin	7.43[1]	7.9 LD[10]		
Edestin	5.94[1]			
Gelatin	2.55[1]	2.2 LA,[7] 2.3 LD[10]		
Human Fibrin	5.0[4]	4.4 LA[7]		
Casein	6.46[1]	5.3 LA,[7] 5.9 LD[10]		
Lactalbumin		3.6 LM[7]		
		Proline		
β-Lactoglobulin		4.1 NL[3]		
Insulin C	2.56[1]			
Insulin V		2.9 LM[7]		
Egg Albumin	3.16[1]	4.1 LM[7]		
Edestin	3.56[1]	4.4 LM[7]		
Gelatin	14.80[1]	15.5 LM[7]		17.5[17]
Human Fibrin	3.85[4]	5.7 LM[7]		
Casein	10.55[1]	11.2 LM[7]		
Lactalbumin		6.0 LM[7]		

Protein	Isolation	Microbiological Assay	Isotope Dilution
		Glutamic Acid	
β-Lactoglobulin	21.51[29]	19.5 LA,[3] 18.7 LA,[40] 18.9 LA[41]	19.1[8]
Insulin C	18.60[2]		
Insulin V		20.2 LA[7]	
Horse Hemoglobin		8.9 LA[41]	8.5[8]
Human Serum Albumin		17.4 LA[7]	17.03[21]
Bovine Serum Albumin		16.9 LA[7]	16.95[21]
Egg Albumin	16.1[30]	15.0 LA,[41] 13.7 LA,[40] 14.3 LA[42]	
Bovine Fibrin		14.3 LA[7]	14.5[19]
Edestin	20.7[30]	21.3 LA,[41] 19.1 LA,[40]	
Gelatin		11.2 LA,[41] 10.8 LA[40]	
Casein	22.01[48]	21.7 LA,[41] 19.7 LA,[40] 21.5 LA,[42] 22.3 LA[7]	
Lactalbumin		15.6 LA[7]	

TABLE 4 (Continued)

Protein	Isolation	Microbiological Assay	Isotope Dilution
		Aspartic Acid	
β-Lactoglobulin	9.9[29]	11.4 LM,[3] 11.5 LM[44]	11.2[8]
Insulin C	5.68[2]		
Insulin V		6.8 LM[7]	
Horse Hemoglobin		10.8 LM[44]	10.3[8]
Human Serum Albumin		10.4 LM[7]	9.88[21]
Bovine Serum Albumin		10.6 LM[7]	10.25[21]
Egg Albumin	8.1[30]	9.3 LM[44]	
Bovine Fibrin		13.0 LM[7]	12.5[19]
Gelatin		6.8 LM[44]	
Casein	6.7[43]	7.2 LM,[44] 7.3 LM[7]	
Lactalbumin		10.9 LM[7]	

Arginine

(1) Microphotometric determination of Brand & Kassell.[24]
(2) Estimated on the catholyte by a modification of the Sakaguchi reaction, Macpherson.[25]
(3) Specific decarboxylase applied to catholyte or to hydrolysate, Gale.[26]
(4) Microbiological assay with *Lactobacilli,* Snell,[5] Guirard & Snell,[27] Stokes *et al.*[10]
(5) Isolation procedures, Chibnall *et al.*,[29, 30] Vickery.[31]
(6) Isotope dilution, Foster.[8]

Protein	(1) Micro. Phot.	(2) Dial. Phot.	(3) Decarb.	(4) MB	(5) Isol.	(6) ID
β-Lactoglobulin	2.87	2.79	2.82	2.91 LC[5] 2.88 SF[27]	2.89[29]	
Insulin C	3.07	3.07			3.04[30]	
Insulin V	3.47					
Horse Hemoglobin		3.63	3.62	3.50 LC[5]	3.62[31] 3.82[29]	3.7
Egg Albumin		5.73		5.6 LC[5] 5.7 SF[27] 5.9 SF[10]	5.66[31] 5.73[29]	
Gelatin		8.51		9.3 LC[5] 9.1 SF[10]	8.56[31]	
Casein		4.09	3.84	3.71 LC[5] 3.71 SF[27] 3.9 SF[10]	3.72[31]	
Lactalbumin						

TABLE 4 (Continued)

Histidine

(1) Colorimetric (Pauli reaction), Hanke & Koessler.[35]
(2) Microphotometric, Pauli reaction on Hg precipitate, Saidel & Brand.[36, 37]
(3) Pauli reaction on catholyte, Macpherson.[25]
(4) Specific decarboxylase applied to catholyte or hydrolysate, Gale.[26]
(5) Microbiological assay with *Lactobacilli,* Brand *et al.,*[7] Stokes *et al.,*[10] Guirard & Snell.[27]
(6) Isolation, 3,4-dichlorobenzene sulfonic acid method of Vickery & Winternitz.[38, 39]

Protein	(1) Color.	(2) Micro. Phot.	(3) Dial. Phot.	(4) Decarb.	(5) MB	(6) Isol.
β-Lactoglobulin		1.58	1.64		1.50 SF[10]	1.54[30]
Insulin C			4.88	4.84		
Insulin V		5.28				
Horse Hemoglobin	8.78	8.79	7.88	7.88 7.75		7.66[35]
Bovine Serum Albumin		3.83			3.80 LM[7]	3.35[38]
Human Serum Albumin		3.52				3.22[38]
Egg Albumin		2.27	2.39	2.46	2.3 SF[10]	
Edestin	3.04	3.02	2.89	2.84		2.52[38]
Gelatin		0.70	0.67		0.58 SF[10] 0.84 LD[27] 0.83 LM[27]	0.57[28]
Casein	2.84	2.88	3.24		2.8 SF[10] 3.2 LD[27] 3.2 LM[27] 2.95 LM[7]	2.52[38]
Lactalbumin					1.9 LM[7]	

TABLE 4 (Continued)

Lysine

(1) Estimated in catholyte by difference, or as picrate, Macpherson.[25]
(2) Specific decarboxylase, applied to catholyte or hydrolysate, Gale,[26] Brand *et al.*[3]
(3) Picrate method, Chibnall *et al.*,[29, 30] Cannan *et al.*,[32] Vickery *et al.*[33]
(4) Microbiological assay with *Lactobacilli*, Brand *et al.*,[7] Stokes *et al.*,[10] Guirard & Snell.[27]
(5) Isotope dilution, Foster,[8] and Shemin.[21]

Protein	(1) Dial. Isol.	(2) Decarb.	(3) Picrate	(4) MB	(5) ID
β-Lactoglobulin	11.1	11.2[26] 11.4[3]	10.55[32]	11.4 LM,[3] 10.4 SF[27] 11.1 SF,[10] 10.5 LM[27]	11.4
Insulin C	2.43	2.47[26]	1.27[30]	2.61 LM[7]	
Horse Hemoglobin	8.2	8.2†[26] 9.2‡[26]	8.1[33]	8.7 SF,[27] 8.45 LM[27]	8.6
Bovine Serum Albumin				12.4 LM[7]	12.42
Egg Albumin	6.08	6.13[26]	5.1[29]	6.6 SF[10]	
Gelatin	4.32*			5.8 SF[10]	
Casein	8.25	8.25[26] 7.5–8.1[34]		7.7 SF,[10] 7.6 SF[27] 7.5 LM[27]	
Lactalbumin				7.7 LM[7]	

* Hydroxylysine has been deducted.
† On catholyte.
‡ On hydrolysate.

Protein	Photometric[46]*	Isotope Dilution
	Tyrosine	
β-Lactoglobulin	3.78[3]	
Insulin V	12.32[7]	
Human Serum Albumin	4.74†[20]	4.73†[21]
Bovine Serum Albumin	5.49[20]	5.53[21]
Gelatin	0.28[7]	
Casein	6.1[7]	
Lactalbumin	3.9[7]	

* Adapted to the Beckman Spectrophotometer.[7]
† Preparation "Comm. I, fraction V."[21]

TABLE 4 (Continued)

Protein	Photometric[46]	Ultraviolet Absorption of Hg Compound[47]
	Tryptophane	
β-Lactoglobulin	1.94	1.94
Human Serum Albumin	0.17	0.19
Bovine Serum Albumin	0.58	0.58
Casein	1.2	1.2
Lactalbumin		1.9
Gelatin	0	0

Protein	Periodate Oxidation	Microbiological Assay
	*Serine**	
β-Lactoglobulin	5.0,[3] 4.1[48]	
Insulin C	5.18[48]	
Insulin V		5.8 LM[7]
Egg Albumin	8.17[48]	10.3 LM[7]
Gelatin	3.1†[48]	
Casein	5.9,[48] 5.8[49]	6.5 LM[7]
	Threonine	
β-Lactoglobulin	5.1,[48] 5.85[3]	4.6 SF[10]
Insulin C	2.08[2]	
Insulin V	3.16[7]	
Egg Albumin	4.0[48]	3.6 SF[10]
Gelatin	2.2[48]	2.0 SF[10]
Casein	4.5[48]	4.2 SF[10]
Lactalbumin	5.1[7]	

* Values are corrected for serine destruction during hydrolysis as follows: 10 per cent[3, 7]; 13 per cent.[48]

† Corrected for hydroxylysine.

Protein	Iodometric[45]	Microbiological Assay[10]
	Methionine	
β-Lactoglobulin	3.22[3]	2.5 SF
Egg Albumin	5.28[45]	4.1 SF
Gelatin	0.90[7]	0.59 SF
Casein	3.1[45]	2.6 SF
Lactalbumin	2.8[45]	

TABLE 4 (Continued)

*Cystine**

Sulfur Partition†

Protein	*Cysteine +* (1)	*Cystine-S* (2)	*Methionine-S* (3)	(2) + (3) (4)	(Pregl) (5)	Protein S (4)/(5) × 100 (6)
	g./100 g.	per cent	per cent	per cent	per cent	per cent
β-Lactoglobulin[3]	3.40	0.905	0.692	1.597	1.604	99.6
Insulin C[2]	12.5‡		0		3.31	100.0
Insulin V	11.6	3.094	0	3.094	3.31	93.5
	12.4§[51]	3.31	0	3.31	3.31	100.0
Chymotrypsinogen[52]	4.50	1.225	0.262	1.487	1.48	100.5
Human Serum Albumin[20]	6.28	1.68	0.275	1.955	1.96	99.7
Bovine Serum Albumin[20]	6.52	1.74	0.18	1.92	1.94	99.0
Casein[45]	0.4	0.09	0.69	0.78	0.78	100.0
Lactalbumin[45]	3.1	0.82	0.60	1.42	1.42	100.0

* Cysteine and cystine determination.[50]
† Sulfur partition.[45]
‡ By calculation.
§ By the Sullivan method (cf.[51]).

method, are the unpublished results of Tristram, quoted by Chibnall in his recent (Sept. 21, 1945) Proctor Lecture.[5] I am greatly indebted to Professor A. C. Chibnall for making the manuscript of his lecture available to me, in advance of publication. Also taken from this lecture,[5] are the unpublished results of Macpherson (arginine, histidine, and lysine estimations on the catholyte), of Rees (modification of the periodate oxidation of serine and threonine), of Gale (various results by specific decarboxylases), and of Sutherland (separate determination of leucine and isoleucine in casein by partition chromatography).

It can be seen from TABLE 4 that, for alanine, we are mainly depending on chromatographic values and for isoleucine on micro bioassays. The chromatographic values for valine and proline are generally lower, sometimes much lower than the corresponding results obtained by micro bioassays. On the other hand, the chromatographic results for phenylalanine are frequently higher than those obtained by bioassays. One should keep in mind, as Snell pointed out,[6] that, if the same value is obtained with two or more organisms, the probability that the results of the micro bioassays are correct is greatly increased.

The micro bioassays for glycine, leucine, glutamic acid, aspartic acid, arginine, and lysine, and the photometric methods for arginine and tyrosine seem to give reliable results, in view of the close agreement with the corresponding values obtained by the isotope dilution and by the solubility product methods.

In the case of histidine, the isolation procedures seem to yield somewhat low results, since the photometric and microbiological methods give uniformly higher figures, which are generally in good agreement with each other. Unfortunately, values obtained by the isotope dilution method are, as yet, not available.

For tryptophane, the results of the photometric determinations are in excellent agreement with the values obtained from the ultraviolet absorption of the tryptophane Hg complex.

The values for serine and threonine, obtained by periodate oxidation, are frequently different from the results of bioassays. This discrepancy is partly due to somewhat arbitrary corrections for destruction during hydrolysis.

Methionine, together with cysteine + cystine, accurately account for the total protein S determined by elementary analysis in the various proteins listed in TABLE 4, insulin being an exception. This would seem to indicate the reliability of the results obtained for the sulfur amino acids. By comparison, the bioassay for methionine yields values that are much too low.

Composition of Proteins

In TABLES 5-A to 5-G, data are given on β-lactoglobulin; in TABLES 6 and 7, on human and bovine serum albumin. TABLE 8 shows a comparison of certain amino acid residues in human and bovine serum albumin. Data on human γ-globulin are presented in TABLES 9-A and 9-B; and preliminary results on the composition of insulin, in TABLES 10-A to 10-E.

β-Lactoglobulin

All the values in TABLE 5-A are our own results; the methods used have been listed in TABLE 1.

Two preparations of β-lactoglobulin were used for the experiments: for one, we are indebted to Dr. R. K. Cannan; for the other, to Dr. G. Haugaard. About 2.5 gm. of protein were used for the determinations reported in TABLE 5-A.

In the interpretation of the data, the following assumptions have been made:

(1) It is assumed that the proteins, subsequently discussed, are chemi-

cal compounds in which the constituent amino acids are present in integral numbers.

(2) It is assumed that the amino acids ($R.NH_2COOH$) are present, in the molecule, in peptide linkage. The yield is, therefore, given in per cent of amino acid and in per cent of amino acid residue R(NH—)(CO—).

TABLE 5-A

COMPOSITION OF β-LACTOGLOBULIN*

(Samples analyzed were prepared by R. K. Cannan and by G. Haugaard.)
$M_{\text{min.}} = 42{,}020$; Total N = 15.60 per cent; Total S = 1.60 per cent

No.	Constituent	gm. of Amino Acid per 100 gm. of Protein	gm. of Amino Acid Residue per 100 gm. of Protein	gm. of Nitrogen per 100 gm. of Protein	Moles per 10^5 gm. of Protein	Residues Moles per Mole	
						(found)	(corr.†)
14	Glycine	1.4	1.06	0.26	18.7	7.8	8
15	Alanine	6.2	4.95	0.97	69.6	29.1	29
16	Valine	5.83	4.93	0.70	49.8	20.9	21
17	Leucine	15.6	13.46	1.67	118.9	50.0	50
18	Isoleucine	8.4	7.25	0.90	64.0	26.9	27
19	Proline	4.1	3.46	0.50	35.6	15.0	15
20	Phenylalanine	3.54	3.15	0.30	21.4	9.0	9
21	Cysteine	1.11	0.94	0.13	9.2	3.9	4
22	Half-Cystine	2.29	1.95	0.27	19.1	8.0	8
23	Methionine	3.22	2.83	0.30	21.6	9.1	9
24	Tryptophane	1.94	1.77	0.27	9.5	4.0	4
25	Arginine	2.88	2.58	0.93	16.5	7.0	7
26	Histidine	1.58	1.40	0.43	10.2	4.3	4
27	Lysine	11.4	10.0	2.18	78.0	32.8	33
28	Aspartic Acid	11.4	9.86	1.20	85.6	36.0	36
29	Glutamic Acid	19.5		1.86			
8a	Amide NH_3	1.31		1.08			
8b	Glutamine‡		9.84		76.8	32.3	32
30	Free Glut. Ac.		7.19		55.7	23.4	24
32	Serine	5.0	4.14	0.67	47.6	20.0	20
33	Threonine	5.85	4.97	0.69	49.1	20.6	21
34	Tyrosine	3.78	3.40	0.29	20.9	8.8	9
37	"Terminal H_2O"		0.20				
38	Total	116.33	99.33	15.60	877.8	368.9	370
39	Water taken up on hydrolysis = 17.00				943.6	396.5	397

* Cf. Brand & Kassell,[24] Brand, Saidel, Goldwater, Kassell, & Ryan.[3]
† Corrected to the nearest integers.
‡ All of amide NH_3 has been arbitrarily assigned to glutamine.

(3) It is assumed that the amide nitrogen or amide NH_3, as obtained by our method, indicates the amount of glutamine and asparagine present. Since it is not yet possible to differentiate between these two amides, all of

the amide NH_3 has been arbitrarily assigned to glutamine and will, subsequently, be discussed as glutamine only. Consequently, all of the aspartic acid is considered as "free aspartic acid," while the "free glutamic acid" is the difference between total glutamic acid and glutamine.

(4) In the interpretation of the data on the free amino nitrogen of the intact protein, obtained in the manometric Van Slyke apparatus under our conditions, it is assumed: (a) that no splitting with the liberation of reactive amino N occurs during the determination; (b) that all of the ε-amino N of lysine is free and reactive; (c) that all of the terminal free α-amino N of the peptide chain, or chains, is free and reactive, and, therefore, can be obtained by deducting the ε-amino nitrogen from the total free amino N; (d) that no proline or hydroxyproline is at the end of the peptide chain; (e) that to each terminal α-amino group corresponds, on the other end of the chain, one free terminal COOH group. In the accounting of the split products in per cent of amino acid residue, a value for "Terminal H_2O" corresponding to the terminal α-amino N has, therefore, been added (*i.e.*, one H atom in terminal NH_2 and one OH group in terminal carboxyl).

(5) It is assumed that our determinations refer to the unhydrated molecule, since they were carried out on air-dried material, corrected for moisture. We have previously contributed some evidence in favor of this contention, in experiments in which H_2O, containing the heavy oxygen isotope O^{18}, was used as an indicator.

Minimum Molecular Weight ($M_{min.}$)

The minimum molecular weight of a protein is defined by the equation:

$$M_{min.} = \frac{R_i \times M_i \times 100}{(\%)_i}, \tag{1}$$

where M_i, $(\%)_i$, and R_i are the molecular weight, the content in per cent and the number of residues of an individual constituent (i). R_i is always an integer. If C_i designates the concentration of the individual constituent in moles per gm., EQUATION 1 can also be written as:

$$M_{min.} = R_i/C_i. \tag{2}$$

If a number of amino acids $(1, 2, 3, \ldots, n)$ has been determined, minimum integral values for $R_1, R_2, R_3, \ldots, R_n$ are obtained by setting up simultaneous equations for $M_{min.}$, yielding a series of equations such as:

$$R_2 = R_1 \times C_2/C_1;\ R_3 = R_1 \times C_3/C_1;\ R_n = R_1 \times C_n/C_1 \tag{3}$$

In these equations, C_1; C_2; C_3 . . . ; C_n are known and, for R_1, a value of 1, 2, 3, 4, etc., is taken successively until integers are found for R_2, R_3 . . ., R_n. In the elementary analysis of organic compounds, the atomic concentrations (C) are usually divided by the smallest individual concentration. With proteins, it is advisable to use as basis a constituent, present in small concentration, which has been determined with great accuracy; and to make the adjustment to integers within 3 per cent of a whole number. This tolerance, imposed by the fact that few amino acids can be determined, at the present time, with an accuracy greater than 3 per cent, is satisfactory, on account of the larger size of the amino acid molecules, as compared to the atomic weights (C, H, N, S, O). With an accuracy of about 3 per cent for the determination of some of the constituent amino acids (present in molar concentrations of 9–35 per 10^5 gm. of protein), satisfactory minimum molecular weights can be obtained in this way for proteins with molecular weights up to about 50,000. For larger minimum molecular weights, a greater accuracy is required. An accuracy of at least 1 per cent for the determination of some constituents would be necessary to obtain significant values for minimum molecular weights of about 175,000. Extensive calculations of minimum molecular weights of proteins from analytical data have been made by Cohn.[7, 8, 9]

In TABLE 5-B, the calculation of the minimum molecular weight is presented. It can be seen that the two constituents present in smallest concentrations are cysteine and tryptophane (9.2 and 9.5 moles per 10^5 gm., respectively). The molecular ratio of these two constituents (within 3 per cent) is 1:1. We have chosen tryptophane as the basis for calculation, because we believe that our tryptophane determination by ultraviolet absorption of the tryptophane mercury complex is particularly accurate. On the other hand, the value for cysteine — although fairly accurate, as indicated from the sulfur partition — is apt to be on the low side, owing to some unavoidable oxidation. Moreover, there is still some controversy about the status of the cysteine residues in the native protein molecule. Serine has been omitted from the calculation, because the value reported in TABLE 5-A includes a somewhat arbitrary correction of 10 per cent for serine destruction during hydrolysis.

In Columns 3 and 4 of TABLE 5-B, the yield in per cent and in moles per 10^5 gm. are given. In Column 5, the molar ratios, on the basis of one tryptophane molecule, are obtained by dividing by the number of moles of tryptophane (9.5). In Columns 6, 7, and 8, these molar ratios, multiplied by 2, 3, and 4, are shown. All figures in Columns 5–8 are rounded out to the nearest integer within 3 per cent. It can be seen that integral ratios, on the basis of 1 or 2 tryptophane residues, are obtained with all amino acids that are present in larger concentrations than 50 moles per 10^5 gm. These constituents, therefore, are not very useful for the establishment of minimum molecular weights. Only 8, 11, and 15 constituents yield integral molar ratios for 1, 2, and 3 tryptophane molecules, respectively, but for 4 tryptophanes there are 19 integral ratios out of a total of 20. The only amino acid that does not yield an integral molar ratio, even for 4 tryptophane molecules, is histidine. It is present in small amount and may not have been determined accurately enough. Histidine is, therefore, omitted from the calculation of the minimum molecular weight (Column 9). This calculation was carried out using EQUATION 1 and the minimum residue numbers in Column 8.

TABLE 5-B

CALCULATION OF MINIMUM MOLECULAR WEIGHT ($M_{min.}$) OF β-LACTOGLOBULIN

(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
No.	Constituent	gm. of Amino Acid per 100 gm. of Protein	Moles per 10^5 gm. of Protein	Molar Ratios, adjusted to Integers within 3 per cent Tryptophane =				$M_{min.}$
24	Tryptophane	1.94	9.5	1	2	3	4	42,110
21	Cysteine	1.11	9.2	1	2	3	4	43,670*
25	Arginine	2.88	16.5	1.74	3.48	5.22	7	42,340
22	Half-Cystine	2.29	19.1	2	4	6	8	41,970
33	Tyrosine	3.78	20.9	2.20	4.40	6.60	9	43,140*
23	Methionine	3.22	21.6	2.27	4.54	7	9	41,700
4	Total S	1.604	50.0	5.26	10.52	16	21	41,970
			Average I (5 Determinations)					42,020 ± 105
14	Glycine	1.4	18.7	2	4	6	8	42,900
20	Phenylalanine	3.54	21.4	2.25	4.50	6.75	9	42,000
19	Proline	4.1	35.6	3.75	7.50	11	15	42,100
32	Threonine	5.85	49.1	5.16	10.32	15.48	21	42,800
16	Valine	5.83	49.8	5.24	10.48	16	21	42,200
30	Glutamic Acid	8.2	55.7	6	12	18	24	43,100
18	Isoleucine	8.4	64.0	6.74	13	20	27	42,200
15	Alanine	6.2	69.6	7.33	15	22	29	41,700
8b	Glutamine	11.2	76.8	8	16	24	32	41,700
27	Lysine	11.4	78.0	8	16	25	33	42,300
28	Aspartic Acid	11.4	85.6	9	18	27	36	42,100
17	Leucine	15.6	118.9	12.5	25	38	50	42,100
			Average II (19 Determinations)					42,300 ± 115
26	Histidine	1.58	10.2	1.07	2.14	3.21	4.28	
Number of Integers (out of 20)				8	11	15	19	

* Values marked are omitted from Average I.

Two average values for $M_{min.}$ are given in TABLE 5-B: 42,020 ± 105 and 42,300 ± 115; the first being based on the five most accurately determined constituents present in small concentrations, the second on all 19. These two values are not significantly different. We shall use the first one (42,020).

Since Pedersen[10] found a molecular weight for β-lactoglobulin in the ultracentrifuge of 41,600, it can be concluded that the minimum molecular weight and the molecular weight are identical and that β-lactoglobulin is monomolecular in solution.

A comparison of values given in the literature for the molecular weight of β-lactoglobulin obtained by different methods is shown in TABLE 5-C.

TABLE 5-C

MOLECULAR WEIGHT OF β-LACTOGLOBULIN*

(1) Minimum by Analysis		$M_{min.}$	42,020 ± 105
Ultracentrifuge†	(2) Sed. vel. and diff.	M_S	41,600
	(3) Sed. equilibrium	M_E	38,000
X-Ray, crystal form‡	(4) Tabular	M_X	[40,000]
	(5) Needle	M_X	[40,000]
(6) Monomolecular Film§		M_F	44,000

* Recently we have noticed that **Gutfreund**[61] has obtained by osmotic pressure measurements a molecular weight, $M_O = 38{,}000$. Since this paper went to press, **Bull & Currie** (J. Am. Chem. Soc. 68: 742. 1946) have reported a new value, $M_O = 35{,}000$.

† **Pedersen.**[66]

‡ Quoted from **Cohn & Edsall;**[62] these values are given in brackets since they refer to "dry" crystals, not corrected for residual water. **McMeekin & Warner**[63] have assumed that the water content of the air-dried crystals used in obtaining the X-ray data was the same as their own air-dried crystals. On this basis and using their own density measurements, they have calculated from the X-ray data an anhydrous molecular weight of 35,800.

§ **Bull.**[64] Since this paper went to press, **Bull & Currie** (J. Am. Chem. Soc. 68: 745. 1946) have reported a new value, $M_F = 2 \times 17{,}000$.

Elementary Composition of β-Lactoglobulin

In order to compare the elementary analysis with the yield of amino acid residues in TABLE 5-A, the percentages of C, H, N, S, and O, corresponding to each constituent, were calculated. The individual percentages thus obtained are omitted from TABLE 5-A, but their sum is given in TABLE 5-D, together with the elementary analysis. The results of the amino acid analysis are somewhat low for oxygen and carbon, the over-all recovery in terms of elements being 99.27 per cent. If the yield of amino acid residues is corrected to the nearest integers, on the basis of a molecular weight of 42,020, the over-all yield in terms of elements is increased to 99.61 per cent.

Sulfur Partition

Methionine, together with cysteine and cystine, account for the total sulfur determined by elementary analysis, sulfate being absent. This concordance constitutes a check on the accuracy of the values for the sulfur amino acids (cf. TABLE 5-E).

TABLE 5-D

ELEMENTARY COMPOSITION OF β-LACTOGLOBULIN

Method	Per Cent of					
	C	H	N	S	O	Total
Elementary Analysis						%
	53.39	7.22	15.60	1.604	22.186	100
Atoms per 42,020	1868	3010	468	21	538	
Amino Acid Analysis						
From the C, H, N, S, O content of the individual amino acid residue.						
A as found	53.09	7.17	15.60	1.597	21.85	99.31
B as corrected to the nearest integers on the basis of $M_{min.}$ = 42,020	53.27	7.23	15.60	1.60	21.93	99.61
Atoms per 42,020	1864	3012	468	21	576	

Nitrogen Partition

TABLE 5-E. All amino acids contain one nitrogen atom in the α position, which we designate the α-N of proteins; this term includes proline and hydroxyproline. Additional N atoms are present in glutamine (amide N), tryptophane (indole N), arginine (guanidino N), histidine (imidazole N), and lysine (ε-amino N). If these amino acids have been determined, their non-α-N can be deducted from the total nitrogen, to yield the α-N. This value, when expressed as atoms of nitrogen per 10^5 gm. of protein (879.4), should be identical with the sum of the number of moles per 10^5 gm. of all the constituent amino acids (877.8), as it has been found to be. In other words, the nitrogen partition derived from the determination of only six constituents agrees closely with the total amino acid composition based on the data for twenty-four constituents. This agreement is further emphasized by a consideration of the amount of water taken up on hydrolysis, 17.00 gm. per 100 gm. of protein. This value, when expressed as moles of H_2O per 10^5 gm. of protein (943.6), should be identical with

the sum of the atoms per 10^5 gm. of the peptide N and of the amide N, if, during hydrolysis, only peptide and amide linkages are split. The calculation, $868.3 + 76.8 = 945.1$ atoms, again shows this to be the case.

TABLE 5-E

SULFUR AND NITROGEN PARTITION OF β-LACTOGLOBULIN

$M_{\min.} = 42{,}020$

No.	Constituent	gm. per 100 gm. of Protein	Atoms per 10^5 gm. of Protein	Atoms per Mole (found)	Atoms per Mole (corr.*)
		Sulfur			
4	Total	1.60	49.9	21.0	21
21a	Cysteine	0.294	9.2	3.9	4
22a	Half-Cystine	0.611	19.1	8.0	8
23a	Methionine	0.692	21.6	9.1	9
6	Protein	1.597	49.8	20.9	21
7	Sulfate	0			0
		Nitrogen			
3	Total	15.60	1113.6	468.0	468
8	Amide	1.076	76.8	32.3	32
24a	Indole	0.133	9.5	4.0	4
25a	Guanidino	0.695	49.6	20.8	21
26a	Imidazole	0.285	20.4	8.6	8
27a	ε-Amino	1.092	78.0	32.8	33
9	Non-α	3.281	234.2	98.4	98
10	α	12.319	879.4	369.5	370
11	Total Amino	1.247	89.0	37.4	37
12	α-Amino	0.155	11.1	4.6	4
13	Peptide	12.164	868.3	365.0	366
	Average Residue Weight = 113.7				
	Average Peptide Weight = 115.2				
	No. of Sub-Units per Mole† = 4				

* Corrected to the nearest integers.
† Corresponding to the number of α-amino N atoms.

The reciprocal of the number of atoms of α-N and of the sum of the moles of the amino acids per 10^5 gm. is the average residue weight (ARW), which is thus found as 113.7 and 113.9, respectively.

The average peptide weight (APW) we find as 114.8, corresponding to 366 peptide bonds per mole. Hotchkiss[11] determined the APW by titration of the increase of amino and carboxyl groups, when β-lactoglobulin is completely hydrolyzed by hydrochloric acid. If his data are recalculated, on the basis of a total N of 15.60 per cent and of a molecular weight of 42,020, a value of 113.0 for APW, corresponding to 372 peptide bonds per mole is obtained, within 2 per cent of our value.

Nature of Ionic Side Chain Groups

A tabulation of these groups is presented in TABLE 5-F. The total number of basic and acid groups found by analysis is in excellent agreement with the titration data of Cannan *et al.*[12]

TABLE 5-F

IONIC SIDE CHAIN GROUPS IN β-LACTOGLOBULIN

	Number per Mole of 42,020 by	
Cationic Groups	Analysis*	Titration†
Guanidino	7	6 (5–7)
Ammonium (ε = 33, α = 4)	37	36 (35–37)
Imidazolyl	4	6
Total	48	48 (46–50)
Anionic Groups		
Carboxyl (Free Dicarb. Acids = 60) (Free Terminal 4)	64	60–63
Total	112	108–111

* Cf. Brand, Saidel, Goldwater, Kassell, & Ryan.[3]

† The data of Cannan, Palmer, & Kibrick[32] were recalculated on the basis of a molecular weight of 42,020.

DISCUSSION

Although practically all of the β-lactoglobulin molecule has been accounted for, we have to consider seriously the possibility that this is the result of numerous compensating errors. There is little doubt that the minimum molecular weight and a considerable number of constituents have been determined with an accuracy that will be subject to little, if any, future change. But is our correction of 10 per cent for destruction of serine during hydrolysis justified? Is our value of 19.5 per cent for glutamic acid, or is Foster's[13] of 19.1 per cent the correct one? Although our value for valine is in excellent agreement with that reported by McMahan and Snell,[14] we have no *independent* evidence that the micro bioassay for this amino acid is completely reliable. Our methods for the determination of proline and of alanine are, certainly, subject to further improvements. Even in the case of leucine, where there is no doubt that both the micro bioassay with the "leucine-less" strain of *Neurospora* and the isotope dilution method give accurate results, there remains the question whether there are 49, 50, or 51 residues of leucine in the molecule. The negative test for hydroxyproline does not exclude the possible presence of a small amount of this amino acid, etc. Some of these difficulties may be resolved in the future by further improvements in the methods of amino acid determinations.

Some conclusions regarding the structure of β-lactoglobulin can be drawn from the data, most of them rather obvious and not particularly startling. We can conclude, with some certainty, that, in this protein and perhaps in many others, the constituent amino acids are primarily linked by typical peptide bonds. Other linkages,* such as esters, anhydrides, or imides, if they are present at all, can be present only in small numbers. There is a possibility that the molecule may contain a small amount of bound water or of water that we have not determined in our usual moisture determination. This is not excluded by the analytical data, since we have not fully accounted for the protein (only 99.62 per cent), nor for all of the oxygen atoms (576, instead of 583). On the other hand, these discrepancies are well within experimental error.

We have interpreted the data for the free α-amino nitrogen as indicating the presence of 4 sub-units or polypeptide chains in the molecule of β-lactoglobulin (based on assumptions no. 4(a) to 4(e)). The only definite conclusion that we can draw, with respect to these sub-units, is that they cannot all be the same, since the minimum molecular weight is identical with the molecular weight. We have no information concerning the size and composition of the sub-units, nor how they are linked chemically in the molecule. To be sure, there are 4 S-S groups in the molecule, which could be imagined to link the sub-units together, but there is, as yet, no evidence that this is actually the case.

The apparently complete analysis of β-lactoglobulin is based upon 28 independent determinations. Of these, 18 were obtained by chemical micro methods; 1 (*lysine*), by a specific enzyme; and 9 (glycine, alanine, valine, *leucine,* isoleucine, proline, phenylalanine, *aspartic* and *glutamic* acid), by microbiological procedures (*italics* indicate close agreement with results obtained by the isotope dilution method[18]).

---

* I am indebted to Dr. John T. Edsall for calling to my attention that, in the paper in the J. Am. Chem. Soc.,[2] it was erroneously concluded that our data excluded the presence of atypical peptide bonds, such as in glutathione. This claim is withdrawn. On the other hand, there is, of course, no evidence that atypical peptide bonds do exist in β-lactoglobulin.

Summary. A summary of the data obtained on β-lactoglobulin is given in TABLE 5-G.

TABLE 5-G

EMPIRICAL FORMULA OF β-LACTOGLOBULIN

Molecular Weight = 42,020
Average Residue Weight = 113.7
Average Peptide Weight = 115.2
370 Amino Acid Residues
4 Sub-Units (Polypeptide Chains)
4 S-S Linkages
Ionic Side Chain Groups = 112 (Cationic = 48; Anionic = 64)
Side Chain Groups capable of forming Hydrogen Bonds = 198

The atomic empirical formula is:

$$C_{1864}\ H_{3012}\ N_{468}\ S_{21}\ O_{576}$$

The empirical formula in terms of amino acid residues (using the first 3 letters of each as the symbol for the amino acid residues) is:

$$\mathrm{Gly_{8}\ Ala_{29}\ Val_{21}\ Leu_{50}\ Ileu_{27}\ Pro_{15}\ Phe_{9}\ (CySH)_{4}\ (CyS\text{-})_{8}\ Met_{9}\ Try_{4}}$$
$$\mathrm{Arg_{7}\ His_{4}\ Lys_{33}\ Asp_{36}\ Glu_{24}\ (Glu\text{-}NH_{2})_{32}\ Ser_{20}\ Thr_{21}\ Tyr_{9}\ (H_{2}O)_{4}}$$

HUMAN AND BOVINE SERUM ALBUMIN

The human serum albumin (TABLES 6-A and 6-B) and bovine serum albumin (TABLES 7-A and 7-B) were crystalline preparations obtained from the Department of Physical Chemistry, Harvard Medical School.[15, 16] The results are given in per cent, in per cent as residue, in per cent nitrogen, in moles per 10^5 gm., and in residues per mole of 70,000. With proteins having the degree of homogeneity possessed by these preparations, the computations, in terms of residues per mole, are of value, because they make the comparison between different proteins easier and may ultimately be of help in the correlation of chemical composition with physico-chemical properties.

TABLE 6-A

COMPOSITION OF HUMAN SERUM ALBUMIN*

Prep. 42

Assumed Molecular Weight = 70,000

No.	Constituent	gm. of Amino Acid per 100 gm. of Protein	gm. of Amino Acid Residue per 100 gm. of Protein	gm. of Nitrogen per 100 gm. of Protein	Moles per 10^5 gm. of Protein	Residues Moles per Mole (found)	(corr.†)
14	Glycine	1.6	1.2	0.30	21.3	14.9	15
16	Valine	7.7	6.5	0.92	65.7	46.0	46
17	Leucine	11.9	10.3	1.27	90.7	63.5	64
18	Isoleucine	1.7	1.5	0.18	13.0	9.1	9
19	Proline	5.1	4.3	0.62	44.3	31.0	31
20	Phenylalanine	7.8	7.0	0.66	47.2	33.0	33
21	Cysteine	0.70	0.60	0.08	5.8	4.1	4
22	Half-Cystine	5.58	4.74	0.65	46.4	32.5	32
23	Methionine	1.28	1.13	0.12	8.6	6.0	6
24	Tryptophane	0.19	0.17	0.03	0.9	0.6	1
25	Arginine	6.15	5.51	1.98	35.3	24.7	25
26	Histidine	3.5	3.1	0.95	22.6	15.9	16
27	Lysine	12.3	10.8	2.36	84.1	58.8	59
28	Aspartic Acid	10.4	9.0	1.09	78.1	54.7	55
29	Glutamic Acid	17.4		1.66			
8a	Amide NH_3	1.07		0.88			
8b	Glutamine‡		8.05		62.8	44.0	44
30	Free Glut. Acid		7.16		55.5	38.9	39
32	Serine	3.7	3.1	0.49	35.2	24.6	25
33	Threonine	5.0	4.2	0.59	42.0	29.4	29
34	Tyrosine	4.66	4.20	0.36	25.7	18.0	18
37	"Terminal H_2O"		0.24				
38	Total	107.7	92.8	15.19	785.2	549.7	551
39	Water taken up on hydrolysis = 14.9				827.0	578.9	579

* Prepared from blood collected by the American Red Cross by the Department of Physical Chemistry, Harvard Medical School, Boston, Mass., under contract, recommended by the Committee on Medical Research, between the Office of Scientific Research and Development and Harvard University.

† Corrected to the nearest integers.

‡ All of the amide NH_3 has been arbitrarily assigned to glutamine.

The data given in TABLES 6 and 7 are more extensive than published previously.[17] Alanine is the only amino acid which has, as yet, not been determined. In the case of the human albumin, about 108 gm. of the split products per 100 gm. of protein are identified and 551 out of 586 amino acids accounted for. For bovine albumin, the data show 112 gm. of the split products and 575 out of 590 amino acids.

TABLE 6-B

NITROGEN PARTITION OF HUMAN SERUM ALBUMIN

Prep. 42

Assumed Molecular Weight = 70,000

No.	Nitrogen	gm. per 100 gm. of Protein	Atoms per 10^5 gm. of Protein	Atoms per Mole
3	Total	15.95	1138.6	797
8	Amide	0.880	62.8	44
24a	Indole	0.013	0.9	1
25a	Guanidino	1.500	107.1	75
26a	Imidazole	0.641	45.4	32
27a	ε-Amino	1.181	84.2	59
9	Non-α	4.215	300.5	210
10	α	11.735	837.7	586
11	Total Amino	1.38	98.5	68
12	α-Amino	0.18	12.8	9
13	Peptide	11.555	824.9	577

Average Residue Weight = 119.4
Average Peptide Weight = 121.2

No. of Sub-Units per Mole* = 9

* Corresponding to the number of α-amino N atoms.

The tryptophane content is 2 residues per mole in bovine, and only about ½ per mole in human, albumin. The fractional value for tryptophane, which has repeatedly been obtained in different highly purified preparations of human albumin, suggests that these preparations are still not chemically homogeneous.

TABLE 7-A

COMPOSITION OF BOVINE SERUM ALBUMIN*

Prep. Armour L-18

Assumed Molecular Weight = 70,000

No.	Constituent	gm. of Amino Acid per 100 gm. of Protein	gm. of Amino Acid Residue per 100 gm. of Protein	gm. of Nitrogen per 100 gm. of Protein	Moles per 10^5 gm. of Protein	Residues Moles per Mole (found)	(corr.†)
14	Glycine	1.9	1.4	0.35	25.3	17.7	18
16	Valine	6.5	5.5	0.78	55.5	38.8	39
17	Leucine	13.7	11.8	1.46	104.5	73.1	73
18	Isoleucine	2.9	2.5	0.31	22.1	15.4	15
19	Proline	5.6	4.7	0.68	48.6	34.1	34
20	Phenylalanine	6.2	5.5	0.53	37.6	26.3	26
21	Cysteine	1.11	0.95	0.12	9.16	6.4	6
22	Half-Cystine	5.41	4.60	0.63	45.0	31.5	32
23	Methionine	0.81	0.71	0.08	5.4	3.8	4
24	Tryptophane	0.58	0.53	0.08	2.84	2.0	2
25	Arginine	6.2	5.6	1.99	35.6	24.9	25
26	Histidine	3.8	3.4	1.03	24.5	17.2	17
27	Lysine	12.4	10.9	2.38	84.8	59.4	59
28	Aspartic Acid	10.6	9.2	1.12	79.6	55.7	56
29	Glutamic Acid	16.9		1.61			
8a	Amide NH_3	1.05		0.86			
8b	Glutamine‡		7.90		61.7	43.2	43
30	Free Glut. Acid		6.87		53.2	37.2	37
32	Serine	4.5	3.7	0.60	42.8	29.9	30
33	Threonine	6.5	5.5	0.76	54.6	38.2	38
34	Tyrosine	5.49	4.94	0.42	30.4	21.3	21
37	"Terminal H_2O"		0.27				
38	Total	112.2	96.5	15.79	823.1	576.1	575
39	Water taken up on hydrolysis = 15.7				871.4	610.0	610

* Obtained from the Department of Physical Chemistry, Harvard Medical School, Boston, Mass.
† Corrected to the nearest integers.
‡ All of amide NH_3 is arbitrarily assigned to glutamine.

TABLE 7-B

NITROGEN PARTITION OF BOVINE SERUM ALBUMIN

Prep. Armour L-18

Assumed Molecular Weight = 70,000

No.	Nitrogen	gm. per 100 gm. of Protein	Atoms per 10^5 gm. of Protein	Atoms per Mole
3	Total	16.07	1147.2	803
8	Amide	0.864	61.7	43
24a	Indole	0.040	2.9	2
25a	Guanidino	1.500	107.1	75
26a	Imidazole	0.681	48.6	34
27a	ε-Amino	1.181	84.3	59
9	Non-α	4.266	304.5	213
10	α	11.804	842.7	590
11	Total Amino	1.38	98.5	69
12	α-Amino	0.20	14.3	10
13	Peptide	11.604	828.4	580

Average Residue Weight = 118.7
Average Peptide Weight = 120.7

No. of Sub-Units per Mole* = 10

* Corresponding to the number of α-amino N atoms.

The hydroxyamino acids (tyrosine, serine, threonine) account for 11.5 per cent of the weight in human, and 14.2 per cent in bovine, albumin. Although high, this is by no means unusual.

Among the monoamino acids, glycine and isoleucine are present in both proteins in relatively small amounts, while the content of valine, proline, phenylalanine, and particularly of leucine is much higher.

There is some difference in the number of residues per mole of cysteine and methionine, but the number of disulfide linkages (half-cystine residues) is the same in both proteins. Very similar is the content of arginine, histidine, lysine,* aspartic acid, glutamic acid, and amide NH_3. The molecules of both contain nearly the same number of nitrogen atoms, of amino acid residues and of sub-units (corresponding to the number of free α-amino nitrogen atoms). A comparison of the similar and dissimilar residues in human and bovine serum albumin is given in TABLE 8.

---

* The values for the lysine content of human and bovine serum albumin, reported in TABLES 6-A and 7-A (12.3 per cent and 12.4 per cent, respectively), were obtained by micro bioassay with *L. mesenteroides* P-60[3] and with the specific decarboxylase of *B. cadaveris*.[18] In the case of the bovine serum albumin, our value is identical with that (12.42 per cent) obtained by Shemin[19] by the isotope dilution method. A discussion of our previously reported[17] low value for the lysine content of serum albumin may be found in Brand *et al.*,[2] footnote no. 42.

TABLE 8

COMPARISON OF HUMAN AND BOVINE SERUM ALBUMIN

Number of Amino Acid Residues or Groups per Mole of 70,000

Constituent	Human	Bovine
Similar		
Total Amino Acids	586	590
Nitrogen Atoms	797	803
Sub-Units*	9	10
Amide NH_3	44	43
Arginine	25	25
Histidine	16	17
Lysine	59	59
Aspartic acid	55	56
Glutamic acid	83	80
Disulfide	16	16
Dissimilar		
Cysteine	4	6
Methionine	6	4
Tryptophane	0–1	2
Glycine	15	18
Leucine	64	73
Isoleucine	9	15
Phenylalanine	33	26
Threonine	29	38
Tyrosine	18	21

* Corresponding to the number of free α-amino N atoms.

Human γ-Globulin

Data on the composition of human γ-globulin are given in TABLES 9-A and 9-B. This sample (no. 36, II–1 (H)) was prepared by the Department of Physical Chemistry of the Harvard Medical School. Fraction II, the analysis of which is here reported, does not represent all of the γ-globulin, but only a portion thereof. A certain amount of high molecular weight material is present in Fraction II, which is, thus, not homogeneous with respect to size. However, it contains extremely little material of very high molecular weight (cf. [16]). The best mean molecular weight for this material may be tentatively taken from the osmotic measurements of Scatchard, Batchelder, and Brown as 156,000.[20]

The results of the amino acid analysis are given in TABLE 9-A, in terms of per cent (gm. of amino acid per 100 gm. of protein), in moles per 10^5 gm., and in residues per mole of 156,000. The data are more extensive than previously reported.[17] As was to be expected from the differences in physico-chemical properties (cf. [15]), the composition of γ-globulin differs markedly from that of the albumin. Compared to serum albumin, human γ-globulin contains fewer basic groups and disulfide linkages. The content of free α-amino nitrogen is only 0.11 per cent, as compared with about 0.2 per cent in the serum albumins. On the other hand, the content of tryptophane and of the hydroxy amino acids (tyrosine, serine, and threonine) is very much higher.

TABLE 9-A

COMPOSITION OF HUMAN γ-GLOBULIN*

Prep. 36, II-1(H)

Assumed Molecular Weight = 156,000

No.	Constituent	gm. of Amino Acid per 100 gm. of Protein	gm. of Amino Acid Residue per 100 gm. of Protein	gm. of Nitrogen per 100 gm. of Protein	Moles per 10^5 gm. of Protein	Residues Moles per Mole (corr.†)
14	Glycine	4.2	3.2	0.78	56.0	87
16	Valine	9.7	8.2	1.16	82.8	129
17	Leucine	9.3	8.0	0.99	70.9	111
18	Isoleucine	2.7	2.3	0.29	20.6	32
19	Proline	8.1	6.8	0.99	70.4	110
20	Phenylalanine	4.6	4.1	0.39	27.9	44
21	Cysteine	0.70	0.60	0.08	5.8	9
22	Half-Cystine	2.37	2.01	0.28	19.7	30
23	Methionine	1.06	0.93	0.10	7.1	11
24	Tryptophane	2.86	2.61	0.39	14.0	22
25	Arginine	4.80	4.30	1.54	27.6	43
26	Histidine	2.50	2.21	0.68	16.1	25
27	Lysine	8.1	7.1	1.55	55.4	86
28	Aspartic Acid	8.8	7.6	0.93	66.1	103
29	Glutamic Acid	11.8		1.12		
8a	Amide NH_3	1.35		1.11		
8b	Glutamine‡		10.15		79.2	124
30	Free Glut. Acid		0.12		1.0	2
32	Serine	11.4	9.5	1.52	108.5	169
33	Threonine	8.4	7.1	0.99	70.5	110
34	Tyrosine	6.75	6.08	0.52	37.3	58
37	"Terminal H_2O"		0.14			
38	Total	109.5	93.1	15.41	836.9	1305
39	Water taken up on hydrolysis = 16.4				910.3	1420

* Prepared from blood collected by the American Red Cross by the Department of Physical Chemistry, Harvard Medical School, Boston, Mass., under contract, recommended by the Committee on Medical Research, between the Office of Scientific Research and Development and Harvard University.

† Corrected to the nearest integers.

‡ All of the amide NH_3 is arbitrarily assigned to glutamine.

Data on the composition of other human plasma protein fractions (α-globulin, β-globulin, fibrinogen, and fibrin), as well as those of some other species, will be published elsewhere. They will permit a direct comparison of the amino acid make-up of plasma proteins, similar to, but in more detail than, the outline previously presented by Brand, Kassell, and Saidel.[17]

TABLE 9-B

NITROGEN PARTITION OF HUMAN γ-GLOBULIN

Prep. 36, II-1(H)

Assumed Molecular Weight = 156,000

No.	Nitrogen	gm. per 100 gm. of Protein	Atoms per 10^5 gm. of Protein	Atoms per Mole
3	Total	16.03	1144.3	1785
8	Amide	1.11	79.2	124
24a	Indole	0.20	14.3	22
25a	Guanidino	1.16	82.8	129
26a	Imidazole	0.45	32.1	50
27a	ε-Amino	0.78	55.7	86
9	Non-α	3.70	264.1	412
10	α	12.33	880.2	1373
11	Total Amino	0.89	63.5	98
12	α-Amino	0.11	7.9	12
13	Peptide	12.22	872.4	1361

Average Residue Weight = 113.6
Average Peptide Weight = 114.6

No. of Sub-Units per Mole* = 12

* Corresponding to the number of α-amino N atoms.

Insulin

In TABLE 10, preliminary data are reported on the composition of a sample of crystalline insulin, for which we are indebted to Dr. V. du Vigneaud. The yield for the various constituents, the calculation of the minimum molecular weight, and the sulfur and nitrogen partition are reported in TABLES 10-A, 10-B, and 10-C, respectively. The preparation contains 3.31 per cent of total sulfur (in close agreement with the value of Miller and du Vigneaud[21]) and 16.04 per cent of total nitrogen (micro-Dumas).

A total of 114.0 gm. of split products per 100 gm. of protein has been obtained, or 97.7 per cent of the protein, on a residue basis. The minimum molecular weight of 44,600, calculated from the values for lysine, arginine, threonine, and histidine, is in good agreement with the value of 46,000 obtained by Miller and Andersson[22] from sedimentation velocity and diffusion data (cf. TABLE 10-B). From TABLE 10-C, it can be seen that only 44 out of 46 sulfur atoms in the molecule have been accounted for: 2 as cysteine and 42 as half-cystine (indicating 21 disulfide bridges). This deficit may, possibly, be due to destruction during hydrolysis. The presence of cysteine (0.6 per cent, determined by our photometric method[23]) was confirmed by a positive Dische test,[24] repeatedly obtained on insulin hydrolysates. This sensitive color reaction with thymonucleic acid is characteristic for sulfhydryl compounds, although not specific for cysteine.

TABLE 10-A

PRELIMINARY DATA ON CRYSTALLINE INSULIN*

$M_{min.} = 44{,}600$; Total N = 16.04 per cent; Total S = 3.31 per cent
(Corrected for 0.8% Ash)

No.	Constituent	gm. of Amino Acid per 100 gm. of Protein	gm. of Amino Acid Residue per 100 gm. of Protein	gm. of Nitrogen per 100 gm. of Protein	Moles per 10^5 gm. of Protein	Residues Moles per Mole (found)	Residues Moles per Mole (corr.†)
14	Glycine	4.6	3.5	0.86	61.3	27.3	27
16	Valine	8.8	7.5	1.05	75.1	33.5	34
17	Leucine‡	13.4	11.6	1.43	102.2	45.6	46
18	Isoleucine	2.9	2.5	0.31	22.1	9.9	10
19	Proline	2.9	2.5	0.35	25.2	11.2	11
20	Phenylalanine	7.9	7.0	0.67	47.8	21.3	21
21	Cysteine	0.6	0.5	0.07	5.0	2.2	2
22	Half-Cystine	11.0	9.4	1.28	91.6	40.9	42
23	Methionine	0		0			0
24	Tryptophane	0		0			0
25	Arginine§	3.5	3.1	1.13	20.1	8.9	9
26	Histidine‖	5.3	4.7	1.44	34.2	15.2	15
27	Lysine¶	2.6	2.3	0.50	17.8	8.0	8
28	Aspartic Acid	6.8	5.9	0.72	51.1	22.8	23
29	Glutamic Acid	20.2		1.92			
8a	Amide NH_3	2.15		1.77			
8b	Glutamine**		16.18		126.2	56.3	56
30	Free Glut. Acid		1.54		11.9	5.3	5
32	Serine	5.8	4.8	0.77	55.2	24.6	25
33	Threonine††	3.2	2.7	0.38	26.9	12.0	12
34	Tyrosine‡‡	12.3	11.1	0.95	67.9	30.3	30
37	"Terminal H_2O"		0.85				
38	Total	114.0	97.7	15.60	841.6	375.3	376
39	Water taken up on hydrolysis = 16.3				904.7	403.5	404

* Obtained from V. du Vigneaud.
† Corrected to the nearest integers.
‡ 13.4 ± 0.3.
§ 3.47 ± 0.03.
‖ 5.30 ± 0.03.
¶ 2.61 ± 0.05.
** All of the amide NH_3 is arbitrarily assigned to glutamine.
†† 3.16 ± 0.03.
‡‡ 12.32 ± 0.03.

The nitrogen partition (reported in TABLE 10-C) indicates 390 amino acid residues in the insulin molecule. There are 369 peptide linkages and 21 sub-units (corresponding to the free α-amino nitrogen). The average residue weight we find to be 114.6 and the average peptide weight, 121.2. This latter value is in good agreement with the APW, obtained by Linderstrøm-Lang,[25] by titration of the increase in amino and carboxyl groups after hydrolysis of insulin. However, this agreement is perhaps less satisfactory than it appears, because of the difference in the values for the total N and for the amide NH_3 of insulin reported by Linderstrøm-Lang and by us. Our values for these constituents are higher than those obtained by other investigators.[5, 25, 26, 27]

TABLE 10-B

CRYSTALLINE INSULIN

Calculation of $M_{min.}$

(Corrected for 0.8% Ash)

No.	Constituent	gm. of Amino Acid per 100 gm. of Protein	Moles per 10^5 gm. of Protein	Molar Ratios, adjusted to Integers within 2 per cent Lysine =					$M_{min.}$
27	Lysine	2.6	17.8	1	2	4	6	8	45,000
25	Arginine	3.5	20.1	1.13	2.26	4.52	7.78	9	44,800
33	Threonine	3.2	26.9	1.51	3	6	9	12	44,600
26	Histidine	5.3	34.2	1.92	3.84	7.68	11.52	15	43,900
						Average			44,600*

* Without ash (mostly Zn) correction = 45,000. M_s = 46,000.[65]

In TABLE 10-A, it is shown that, on a residue basis, 97.7 per cent of the protein and 376 amino acid residues are accounted for. A calculation is presented in TABLE 10-D, which indicates that the N content and the residue weight of the undetermined residues are similar to those of alanine (corresponding to a yield of 2.9 gm. of alanine per 100 gm. of insulin).

The elementary composition of insulin is shown in TABLE 10-E. The elementary analysis is in good agreement with the C, H, N, S, and O content, calculated from the elementary content of the individual amino acid residues reported in TABLE 10-A. This agreement is particularly significant in the case of oxygen. Elementary analysis indicates 583 oxygen atoms per mole, while 590 oxygen atoms per mole are calculated from the analytical data for the individual constituents. The latter figure does not involve any assumption regarding the nature of the 14 undetermined amino acid residues.

A condensed summary of the data on insulin is given in TABLE 10-F.

In Chibnall's Proctor Lecture,[5] an almost complete analysis of crystalline insulin is given. All the data were obtained in the Biochemical Laboratory at Cambridge. The total nitrogen and the amide NH_3 of the insulin analyzed by the English workers are 15.54 and 1.68 per cent, respectively, as compared to 16.04 and 2.15 per cent in TABLES 10-A and 10-E. The total (S) of Chibnall's preparation is the same as that reported by Miller and

TABLE 10-C

SULFUR AND NITROGEN PARTITION OF CRYSTALLINE INSULIN

$M_{min.} = 44,600$*

(Corrected for 0.8% Ash)

No.	Constituent	gm. per 100 gm. of Protein	Atoms per 10^5 gm. of Protein	Atoms per Mole (found)	Atoms per Mole (corr.†)
		Sulfur			
4	Total	3.31	103.2	46.0	46
21a	Cysteine	0.159	5.0	2.2	2
22a	Half-Cystine	2.935	91.6	40.9	42‡
23a	Methionine	0			0
6	Protein	3.094	96.6	43.1	44
7	Sulfate	0			0
	Not Accounted			2.9	2
		Nitrogen			
3	Total	16.04§	1145.1	510.7	511
8	Amide	1.768	126.2	56.2	56
24a	Indole	0			0
25a	Guanidine	0.844	60.3	26.8	27
26a	Imidazole	0.958	68.4	30.4	30
27a	ε-Amino	0.249	17.8	7.9	8
9	Non-α	3.819	272.6	121.4	121
10	α	12.22	872.4	389.1	390
11	Total Amino	0.91	65.0	29.0	29
12	α-Amino	0.66	47.1	21.0	21
13	Peptide	11.56	825.3	368.1	369

Average Residue Weight = 114.6
Average Peptide Weight = 121.2

No. of Sub-Units per Mole‖ = 21

* $M_S = 46,000$.[65]

† Corrected to the nearest integers.

‡ By definition, half-cystine residues must be an even number per mole.

§ 16.04 ± 0.04.

‖ Corresponding to the number of α-amino N atoms.

du Vigneaud and by us, 3.31 per cent[21] (TABLE 10-C). However, cystine was not determined, but assumed to account for all of the sulfur in insulin.

Table 10-D

Summary of Preliminary Data on Crystalline Insulin

$M_{min.} = 44,600$ Amino Residues / Mole = 390 Total N = 16.04 per cent
Peptide Linkages per Mole = 369 Amide Linkages per Mole = 56
Water taken up on hydrolysis per Mole expected: 369 + 56 = 425 Moles

Accounted:	
97.7 per cent of Protein (as Residue)	376 Amino Acid Residues
	15.60 per cent N
Not Accounted: (Alanine not determined)	14 Amino Acid Residues
2.3 per cent of 44,600 = 1030	0.44 per cent N

Undetermined Amino Acid Residues

Average Residue Weight = 1030 / 14 = 73 (RW of Alanine = 71)
N Content = 0.44 / 2.3 × 100 = 19.1 per cent (Alanine Residue contains 19.7 per cent N)

Water of Hydrolysis	*Moles per mole*
Determined	404
From undetermined Amino Acid Residues	14
Total	418
Expected from Peptide and Amide Linkages	425

Table 10-E

Elementary Composition of Insulin

$M_{min.} = 44,600$

Elementary Analysis	C	H	N	S	O	Total
Per cent	52.96	6.79	16.04	3.31	20.91	100
Atoms/mole	1966	3004	511	46	583	
Amino Acid Analysis						
from C, H, N, S, O content of the individual amino acid residues in TABLE 10-A (alanine from TABLE 10-D).						
Per cent	52.84	6.64	16.05	3.09	21.16	99.78
Atoms/mole	1962	2939	511	44	590*	

* This figure does not involve any assumptions regarding the nature of the 14 undetermined amino acid residues.

For tyrosine, Chibnall reports a yield of 13.03 gm. per 100 gm., while we find only 12.32 gm. For the other amino acids, our yield is uniformly higher (Chibnall's values are given in TABLE 4, Insulin C).

Of fundamental importance is a new method, developed by F. Sanger,[28] in Chibnall's laboratory, for the identification and estimation of the free amino groups in proteins and peptides, using 2,4-dinitrofluorobenzene. Sanger[28] (cf. [5]) has applied this method to insulin and found that glycine and phenylalanine occupy terminal positions in the peptide chains that constitute the insulin molecule (qualitative evidence that phenylalanine is a terminal amino acid had previously been obtained by Jensen and Evans[29]).

Chibnall's data lead him to accept a minimum molecular weight of 12,000 per submolecule of insulin. He concludes that the insulin molecule is an association of either 3 or 4 such submolecules, depending upon whether a molecular weight, $M_X = 36{,}000$ (from X-ray data of Crowfoot[30]), or the much higher value, $M_S = 46{,}000$, of Miller and Andersson,[22] is accepted.

TABLE 10-F

CRYSTALLINE INSULIN

$M_{min.} = 44{,}600$ 390 Residues 21 Sub-Units

Tentative Empirical Formula in Terms of Amino Acid Residues

The first 3 letters of the amino acids are used as symbols for the corresponding residues.* (XSR) indicates an undetermined sulfur containing residue; (XAR), an undetermined amino acid residue.

$$\mathrm{Gly_{27}\ Val_{34}\ Leu_{46}\ Ileu_{10}\ Pro_{11}\ Phe_{21}\ (CySH)_{2}\ (CyS\text{-})_{42}\ Arg_{9}\ His_{15}\ Lys_{8}}$$
$$\mathrm{Asp_{23}\ Glu_{5}\ (Glu\text{-}NH_2)_{56}\ Ser_{25}\ Thr_{12}\ Tyr_{30}\ (XSR)_{2}\ (XAR)_{12}\ (H_2O)_{21}}$$

The (XSR) residues are possibly cystine; most of the (XAR) residues are probably alanine.

* **Cf. Brand, Saidel, Goldwater, Kassell, & Ryan.[3]**

In the case of insulin, the discrepancy between M_X and M_S is even larger than it is in the case of β-lactoglobulin ($M_X = 36{,}000$, $M_S = 42{,}000$, cf. TABLE 5-C). For both proteins, $M_{min.}$, calculated from our analytical data, agrees with the higher values obtained from sedimentation velocity and diffusion data. This agreement is significant only if our analyses are accurate within the limits claimed: about 3 per cent for β-lactoglobulin (TABLE 5-C) and 2 per cent for insulin (TABLE 10-B).

Proposed Abbreviations and Symbols

These are summarized in TABLE 11. It will become imperative, in the future, to find a compact way of presenting the empirical formula of a protein in terms of its amino acid residues. We propose to use as symbols, whenever feasible, the first 3 letters of the amino acids: *e.g.*, glycine residue, Gly.; arginine residue, Arg.; isoleucine residue, Ileu., etc. The residues of cysteine, half-cystine, asparagine and glutamine we propose to designate by CySH, (CyS-), (Asp-NH_2), and (Glu-NH_2), respectively. Such formulae for β-lactoglobulin and for insulin are given in TABLES 5-G and 10-F, respectively. I hope that a satisfactory convention will be gradually established.

TABLE 11

PROPOSED TERMINOLOGY AND ABBREVIATIONS

Molecular Weight

$M_{min.}$ = minimum from analysis.
M_S = from sedimentation velocity and diffusion.
M_E = from sedimentation equilibrium.
M_O = from osmotic pressure.
M_X = from X-ray data.
M_F = from monomolecular films.

α-N = α Nitrogen of the constituent amino acids (not α-amino N).
"Terminal H_2O" = Water in terminal NH_2 and COOH groups, not in peptide linkage.
Water taken up on hydrolysis = yield in per cent of amino acid less yield in per cent of amino acid residue.
Sub-Units = Corresponding to the number of free α-Amino N atoms.

ARW = Average Residue Weight.
APW = Average Peptide Weight.

Symbols for Amino Acid Residues

Gly
Ala
Val
Nval
Leu
Ileu
Pro
Phe
(CySH)
(CyS-)
Met
Try
Arg
His
Lys
Asp
Glu
(Glu-NH_2)
(Asp-NH_2)
Ser
Thr
Tyr
Hypro
Hylys

SUMMARY

Micro methods, developed in this laboratory for the determination of protein constituents, are briefly reviewed. They include ultraviolet absorption spectroscopy; micro-diffusion, titration, and photometric and enzymatic techniques; elementary analysis and bioassays with mutants of *Neurospora* and different strains of *Lactobacilli*.

A comparison of results, obtained by our methods and by others, is given for a number of proteins. Included in the tabulation are the following types of methods: Isotope Dilution, Solubility Product, Partition

Chromatography, Isolation, Photometry, Oxidation, Iodometry, Specific Decarboxylases, and Bioassays.

Data on the composition of β-lactoglobulin, bovine serum albumin, human serum albumin, human γ-globulin, and crystalline insulin are shown. Alanine is the only amino acid which has not been determined in these proteins (except in β-lactoglobulin).

For β-lactoglobulin and for insulin, the calculation of the minimum molecular weights and condensed summaries of the empirical formulae are given.

Symbols to designate amino acid residues are proposed.

BIBLIOGRAPHY

1. **Mulder, J. G.**
   1838. Ann. **28,** 73.
2. **Brand, E., L. J. Saidel, W. H. Goldwater, W. B. Kassell, & F. J. Ryan**
   1945. J. Am. Chem. Soc. **67:** 1524.
3. **Brand, E., & L. J. Saidel**
   Unpublished experiments.
4. **Martin, A. J. P., & R. L. M. Synge**
   1941. Biochem. J. **35:** 1358.
5. **Chibnall, A. C.**
   1946. J. Int. Soc. Leather Trades Chem. **30:** 1.
6. **Snell, E. E.**
   1946. Ann. N.Y. Acad. Sci. **47** (2): 161–179.
7. **Cohn, E. J., J. L. Hendry, & A. M. Prentiss**
   1925. J. Biol. Chem. **63:** 721.
8. **Cohn, E. J.**
   1939. Chem. Rev. **24:** 203.
9. **Cohn, E. J.**
   1939. Bull. N.Y. Acad. Med. **15**(II): 639.
10. **Pedersen, K. O.**
    1936. Biochem. J. **30:** 948, 961.
11. **Hotchkiss, R. D.**
    1939. J. Biol. Chem. **131:** 387.
12. **Cannan, R. K., A. H. Palmer, & A. C. Kibrick**
    1942. J. Biol. Chem. **142:** 803.
13. **Foster, G. L.**
    1945. J. Biol. Chem. **159:** 431.
14. **McMahan, J. R., & E. E. Snell**
    1944. J. Biol. Chem. **152:** 83.
15. **Cohn, E. J., J. L. Oncley, L. E. Strong, W. L. Hughes, Jr., & S. H. Armstrong, Jr.**
    1944. J. Clin. Invest. **23:** 417.
16. **Williams, J. W., M. L. Petermann, G. C. Colovas, M. B. Goodloe, J. L. Oncley, & S. H. Armstrong**
    1944. J. Clin. Invest. **23:** 433.
17. **Brand, E., B. Kassell, & L. J. Saidel**
    1944. J. Clin. Invest. **23:** 437.

18. **Gale, E. F.**
1945. Biochem. J. **39:** 46.
19. **Shemin, D.**
1945. J. Biol. Chem. **159:** 439.
20. **Scatchard, G., A. C. Batchelder, & A. Brown**
1944. J. Clin. Invest. **23:** 458.
21. **Miller, G. L., & V. du Vigneaud**
1937. J. Biol. Chem. **118:** 101.
22. **Miller, G. L., & K. J. I. Andersson**
1942. J. Biol. Chem. **144:** 459.
23. **Kassell, B., & E. Brand**
1938. J. Biol. Chem. **125:** 115.
24. **Dische, L.**
1944. Proc. Soc. Exp. Biol. & Med. **55:** 217.
25. **Linderstrøm-Lang, K.**
1940. Enzymologia **8:** 64.
26. **Harington, C. R., & T. H. Mead**
1936. Biochem. J. **30:** 1598.
27. **Chibnall, A. C.**
1942. Proc. Roy. Soc. London B **131:** 136.
28. **Sanger, F.**
1945. Biochem J. **39:** 507.
29. **Jensen, H. F., & E. A. Evans, Jr.**
1935. J. Biol. Chem. **108:** 1.
30. **Crowfoot, D.**
1938. Proc. Roy. Soc. London A **164:** 580.

BIBLIOGRAPHY FOR TABLES

1. **Tristram, G. R.**
Unpublished. Quoted by Chibnall.[2]
2. **Chibnall, A. C.**
1946. J. Int. Soc. Leather Trades Chem. **30:** 1.
3. **Brand, E., L. J. Saidel, W. H. Goldwater, B. Kassell, & F. J. Ryan**
1945. J. Am. Chem. Soc. **67:** 1524.
4. **Tristram, G. R.**
Unpublished. Quoted by K. Bailey.[4a]
4a. **Bailey, K.**
1944. Advances in Protein Chemistry **1:** 289. Academic Press. New York.
5. **McMahan, J. R., & E. E. Snell**
1944. J. Biol. Chem. **152:** 83.
6. **Kuiken, K. A., W. H. Norman, C. M. Lyman, F. Hale, & L. Blotter**
1943. J. Biol. Chem. **151:** 615.
7. **Brand, E., & L. J. Saidel**
Unpublished data.
8. **Foster, G. L.**
1945. J. Biol. Chem. **159:** 431.
9. **Sutherland**
Quoted by Chibnall.[2]
10. **Stokes, J. L., M. Gunness, I. M. Dwyer, & M. Caswell**
1945. J. Biol. Chem. **160:** 35.
11. **Brand, E., F. J. Ryan, & E. M. Diskant**
1945. J. Am. Chem. Soc. **67:** 1532.

12. **Ryan, F. J., & E. Brand**
1944. J. Biol. Chem. **154**: 161.

13. **Moore, S., & W. H. Stein**
1943. J. Biol. Chem. **150**: 113.

14. **Schweigert, B. S., J. M. McIntire, C. A. Elvehjem, & F. M. Strong**
1944. J. Biol. Chem. **155**: 183.

15. **Stein, W. H.**
1942. Am. Soc. Brewing Chem. Proc. 5th Ann. Meeting **13.**

16. **Bergmann, M., & W. H. Stein**
1939. J. Biol. Chem. **128**: 217.

17. **Stein, W. H., & M. Bergmann**
1940. J. Biol. Chem. **134**: 627.

18. **Sanger, F.**
1945. Biochem. J. **39**: 507.

19. **Rittenberg, D., & G. L. Foster**
1940. J. Biol. Chem. **133**: 737. As quoted by Brand *et al.*[20]

20. **Brand, E., B. Kassell, & L. J. Saidel**
1944. J. Clin. Invest. **23**: 437.

21. **Shemin, D.**
1945. J. Biol. Chem. **159**: 439.

22. **Haugaard, G.**
Unpublished. Quoted by Stein and Moore.[23]

23. **Stein, W. H., & S. Moore**
1946. Ann. N.Y. Acad. Sci. **47** (2): 95–118.

24. **Brand, E., & B. Kassell**
1942. J. Biol. Chem. **145**: 359.

25. **Macpherson, H. T.**
Unpublished. Quoted by Chibnall.[2]

26. **Gale, E. F.**
1945. Biochem. J. **39**: 46; and unpublished. Quoted by Chibnall.[2]

27. **Guirard, B. M., & E. E. Snell**
Unpublished. Quoted by Snell.[28]

28. **Snell, E. E.**
Ann. N.Y. Acad. Sci. **47** (2): 161–179

29. **Chibnall, A. C., M. W. Rees, & E. F. Williams**
1943. Biochem. J. **37**: 372.

30. **Chibnall, A. C.**
1942. Proc. Roy. Soc. London B **131**: 136.

31. **Vickery, H. B.**
1940. J. Biol. Chem. **132**: 325.

32. **Cannan, R. K., A. H. Palmer, & A. C. Kibrick**
1942. J. Biol. Chem. **142**: 803.

33. **Vickery, H. B., & C. F. Leavenworth**
1928. J. Biol. Chem. **79**: 377.

34. **Zittle, C. A., & V. R. Eldred**
1944. J. Biol. Chem. **156**: 401.

35. **Hanke, M. T., & K. K. Koessler**
1920. J. Biol. Chem. **43**: 527.

36. **Saidel, L. J., & E. Brand**
1944. (September). Am. Chem. Soc. (New York Meeting). Abstracts: B-26.

37. **Saidel, L. J., & E. Brand**
Unpublished experiments.

38. **Vickery, H. B., & J. K. Winternitz**
1944. J. Biol. Chem. **156**: 211.
39. **Vickery, H. B.**
1944. J. Biol. Chem. **156**: 283.
40. **Lewis, J. C., & H. S. Olcott**
1945. J. Biol. Chem. **157**: 265.
41. **Hac, L. R., E. E. Snell, & R. J. Williams**
1945. J. Biol. Chem. **159**: 273.
42. **Lyman, C. M., K. A. Kuiken, L. Blotter, & F. Hale**
1945. J. Biol. Chem. **157**: 395.
43. **Bailey, K., A. C. Chibnall, M. W. Rees, & E. F. Williams**
1943. Biochem. J. **37**: 360.
44. **Hac, L. R., & E. E. Snell**
1945. J. Biol. Chem. **159**: 291.
45. **Kassell, B., & E. Brand**
1938. J. Biol. Chem. **125**: 145, 435.
46. **Brand, E., & B. Kassell**
1939. J. Biol. Chem. **131**: 489.
47. **Brand, E., & L. J. Saidel**
1943 (September). Am. Chem. Soc. (Pittsburgh Meeting). Abstracts: B-22.
48. **Rees, M. W.**
Unpublished. Quoted by Chibnall.[2]
49. **Nicolet, B. H., & L. A. Shinn**
1941. J. Biol. Chem. **139**: 687.
1942. Ibid. **142**: 139.
50. **Kassell, B., & E. Brand**
1938. J. Biol. Chem. **125**: 115, 131.
51. **Miller, G. L., & V. du Vigneaud**
1937. J. Biol. Chem. **118**: 101.
52. **Brand, E., & B. Kassell**
1941. J. Gen. Physiol. **25**: 167.
53. **Saschek, W.**
1937. Ind. & Eng. Chem. Anal. Ed. **9**: 491.
54. **Saschek, W., & E. Brand**
Unpublished experiments.
55. **Saidel, L. J., & E. Brand**
Unpublished experiments.
56. **Warner, R. C., & R. K. Cannan**
1942. J. Biol. Chem. **142**: 725.
57. **Diskant, E. M., & E. Brand**
Unpublished.
58. **Shinn, L. A., & B. H. Nicolet**
1941. J. Biol. Chem. **138**: 91.
59. **Winnick, T.**
1942. J. Biol. Chem. **142**: 461.
60. **McFarlane, W. D., & G. N. Guest**
1939. Canad. J. Res. **17**: 139.
61. **Gutfreund, H.**
1945. Nature **155**: 237.
62. **Cohn, E. J., & J. T. Edsall**
1943. Proteins, Amino Acids and Peptides: 328. Reinhold Publishing Corp. New York.

63. **McMeekin, T. L., & R. C. Warner**
1942. J. Am. Chem. Soc. **64**: 2393.
64. **Bull, H. B.**
1945. J. Am. Chem. Soc. **67**: 8.
65. **Miller, G. L., & K. J. I. Andersson**
1942. J. Biol. Chem. **144**: 459.
66. **Pedersen, K. O.**
1936. Biochem. J. **30**: 961.

SOME CORRELATIONS BETWEEN PHYSICO-CHEMICAL DATA AND THE AMINO ACID COMPOSITION OF SIMPLE PROTEINS

Discussion of the Paper by Erwin Brand

By John T. Edsall

Department of Physical Chemistry, Harvard Medical School,
Boston, Massachusetts

The data presented in Dr. Brand's paper appear to indicate the beginning of a new epoch in our knowledge of protein composition. Doubtless, some of these figures will need a moderate amount of revision, and there are a few gaps, here and there. However, we are now beginning, for the first time, to have assured, detailed, and practically complete knowledge of protein composition. It is of particular importance that all the analyses can now be carried out on the same sample of protein, and only a relatively small sample is required.

Two major advances in recent years, both already discussed in detail at this Conference, have been outstanding in bringing this about: First, the introduction of the isotope dilution method, which has given a reliable basis for the accurate estimation of any amino acid, and is clearly the ultimate standard of reference for all such estimations at the present time; second, the introduction of the microbiological methods, which have furnished such rapid and astonishingly accurate estimations of small amounts of many amino acids, including several that were previously quite inaccessible to accurate estimation. The future significance of partition chromatography and allied techniques may be very great indeed, but, at present, the figures derived from these methods appear to be subject to a somewhat greater error than those derived from some of the other methods under discussion. It is probable, as Martin and Synge have suggested elsewhere, that the chromatographic techniques may find their chief application in the separation of closely related peptides, among the products of partial hydrolysis of proteins.

Some use has been made, in the past, of amino acid determinations as a guide to the process of fractionation of a complex mixture of proteins. With the aid of some of these newer methods, we may expect that analytical determinations, as a guide to protein fractionation, will be used far more frequently and to better effect, since the methods, in general, are now more accurate and they require smaller quantities of material. Since the microbiological methods are highly specific for such closely related residues as leucine and isoleucine, all these individual amino acids have become almost equally available as analytical tools in guiding fractionation. Moreover, it may be expected that analytical determinations of amino acid residues will be used far more often in the future as a criterion of purity in protein preparations. Obviously, in a solubility experiment, the percentage of every amino acid residue in the protein of the solid phase must be the same as that of the protein in the supernatant liquid, if the protein is a chemical individual. (Compare the discussion by Shedlovsky.[1])

The estimates of minimum molecular weight from analytical data clearly call for critical comparison with the values obtained by various methods of physical chemistry. A few years ago, the general agreement of estimates derived from these different sources was noted and the apparent discrepancies were treated as minor matters. Now, the techniques available are sufficiently precise to justify emphasizing the discrepancies rather than the agreements. These discrepancies are often relatively minor, but their careful study may well lead the way to new discoveries.

In the first place, consider the molecular weights, obtained from sedimentation and diffusion data, according to the fundamental equation:

$$M = \frac{RTs}{D(1 - V\rho)}. \tag{1}$$

The three variables in this equation which require the most critical examination are the sedimentation constant, s, the diffusion constant, D, and the partial specific volume, V.* The sedimentation constants, determined by different investigators on different samples of the same purified protein, are commonly in very satisfactory agreement. On the other hand, there are frequently much larger disagreements between the diffusion constant values reported by different authors. Thus, the sedimentation constant, s_{20}, reported for insulin in the recent study by Miller and Andersson,[5] was 3.55 S, which differs by less than 3 per cent from the earlier value of 3.47 given by Sjögren and Svedberg.[6] However, the diffusion constant, D_{20}, as found by Miller and Andersson, was 7.53×10^{-7}; whereas the earlier value of Polson,[7] determined by essentially the same method, was 8.2×10^{-7}.

Another example of a striking disagreement between the results obtained by different authors occurs in the diffusion constant measurements for the largest component of *Helix* hemocyanin. Sedimentation constants have been reported for this molecule in several investigations from the Upsala laboratories. All but the earliest of these determinations agree within very narrow limits, giving $s_{20} = 101 \pm 3S$. On the other hand, the diffusion constant, as reported by Polson,[7] differs by about 30 per cent from that reported by Brohult,[8] although both determinations were carried out in the same laboratory, by essentially the same methods. In this case, the great size of the molecule makes diffusion slow and the errors are probably somewhat greater than would be the case for smaller molecules.

---

* Naturally, the present discussion makes no attempt to cover the field adequately. The most complete discussion of the derivation of EQUATION 1 is given by **Svedberg & Pedersen.**[2] Important, critical discussions of the molecular weight values derived from EQUATION 1 have been given, at previous meetings of this Section of the New York Academy of Sciences, by **Oncley**[3] and **Rothen.**[4]

There are several reasons for the fact that the probable errors in the diffusion constants are considerably larger than for the sedimentation constants. These cannot be discussed here in detail. However, it should be remembered that the process of diffusion is very slow and requires several days, whereas a sedimentation run is generally completed within a few hours. During the long time required for a diffusion experiment, partial denaturation of the protein may occur. The difficulties of forming a sharp initial boundary are also important as a disturbing factor. It is interesting to note that the earlier measurements in the literature on any given protein almost always led to higher values of the estimated diffusion constant and, therefore, to lower molecular weights than those which were obtained in the later and, presumably, more accurate studies.

A word should also be said about the partial specific volume, V. The importance of this factor has not always been sufficiently appreciated by those who have made use of the results of ultracentrifuge studies. In EQUATION 1, ρ, the density of the solvent, is generally near unity, whereas V is commonly near 0.75. Hence, an error of 1 per cent in V produces an error of about 3 per cent in the factor $(1 - V\rho)$, and hence, again, in the molecular weight. The partial specific volumes of most proteins hitherto studied are fairly close to 0.75, but the value for gelatin is only 0.68, while that of zein is 0.77, as a recent study from our laboratory, by Foster and French,[9] has shown. The earlier tentative value for zein, as listed by Svedberg and Pedersen,[2] was 0.71. The study of the specific volume of this prolamine in ethanol-water was a particularly difficult technical problem, which was very ingeniously solved by Foster and French, using a method that should be applicable to a number of other proteins. It is worth noting that, when their partial specific volume is used instead of the earlier value, the estimated molecular weight of zein is changed from 38,000 to 50,000, using identical values for s and D, in both cases.

These differences in partial specific volume of different proteins are systematically related to their composition (Cohn and Edsall,[10] Chapter 16).

The low value of V for gelatin is associated with its high glycine content, the glycyl residue being very compact; while the high value of V for zein is correlated with the large numbers of leucyl and isoleucyl residues which it contains, the latter being of relatively low density.

Molecular weights, determined by the sedimentation equilibrium method, should, in principle, be highly reliable. In practice, they are generally regarded with more mistrust than those obtained by sedimentation and diffusion. This is primarily because of the very long time required for true equilibrium to be obtained, a fact which has two undesirable consequences: (1) There is, generally, a slight, remaining doubt whether the observed distribution, even after many days, really does represent a true equilibrium distribution; (2) during these long runs, the protein may become wholly or partially denatured. Archibald[11] has made calculations of the distribution of protein molecules to be expected in the ultracentrifuge cell, when equilibrium is being approached, but has not been fully attained. So far as I am aware, however, his calculations have not yet been applied to actual molecular weight determinations.

X-ray diffraction measurements give, directly, only the volume of the unit cell in a crystal. To obtain a molecular weight, two further quantities must be known: the density of the crystal, and the number of molecules in the unit cell. The latter figure must be an integer, generally small, and can, generally, be estimated without difficulty, if other values of the molecular weight of the same protein, even very rough approximations, are available. Accurate figures for the density of protein crystals, however, are very difficult to get. This is the chief obstacle, at present, to the attempt to use the X-ray data to give accurate molecular weights (Fankuchen[12]). Thus, the value of the density of dry β-lactoglobulin, as given by McMeekin and Warner,[13] when combined with the X-ray measurements of Crowfoot,[14] leads to an estimated molecular weight of only 35,000 — decidedly lower than the value derived from analysis or from the ultracentrifuge studies. This serious discrepancy still remains to be explained.

One of the most impressive results of the work of Snell, Brand, and others who have used microbiological assays is the increasingly precise knowledge of the different amino acids with non-polar side-chains. The physical chemists, like myself, have commonly stressed the importance of the polar, and, especially, the ionic, side-chains. In so far as we considered the non-polar side-chains, our tendency has been to lump them together as a group; or to think in terms of the total numbers of alkyl groups or of aromatic rings in the molecule, rather than in terms of the number and configuration of specific non-polar groups. Their precise structural and biological significance in the protein molecule is still far from clear. However, I am convinced that they have profoundly important and highly specific functions and that such side-chains as those of valine, leucine, and isoleucine are definitely required to fit into certain portions of the protein molecule. They are certainly not interchangeable. The specific need for valine, leucine, and isoleucine in nutrition is only one indication of this. For purely steric reasons, one can already note some basis for specificity of this sort.

The amino acids with acidic and basic side-chains, however, have long been susceptible to independent measurement, both by the analytical chemist and the physical chemist. Dr. Brand has already given a brief tabular comparison between his data on β-lactoglobulin and the titration data of Cannan, Palmer, and Kibrick.[15] The comparison in this case is beautifully satisfactory; nothing better could be asked. I should like to offer a more tentative comparison between Brand's data on human and bovine serum albumin, and insulin, with some titration data from our own laboratory.* These data do not represent as complete and detailed titrations as that of β-lactoglobulin, but they permit some comparisons which may be instructive.

In TABLE 1, the data are listed in terms of the number of groups or residues per 100 kg. of protein, in order to emphasize the fact that the relations involved are purely stoichiometric and quite independent of the assumed molecular weights.

Clearly, unless there are unknown ionic groups not yet determined by analysis, the maximum positive charge on the protein, as determined by titration from the isoionic point to a pH between 1 and 2, should correspond to the sum of the histidine, arginine, lysine, and free α-amino groups. Actually, the figures from analysis (Row e in TABLE 1) are higher in all three cases than those from titration (Row f). For both types of serum albumin, the titration figures agree quite closely with the sum of the histidine, arginine, and lysine. If one had only these and the titration figures to go on, it would be natural to assume that there were no free α-amino (and, hence, no α-carboxyl) groups in these molecules, within the limits of experimental error. The situation in insulin is not so simple. The sum of the histidine, arginine, and lysine values is only 72.1, while the titration figure is 100. In this case, one would be led to infer some free α-amino groups from the titration data, although the number is only 12 or 13 per molecule,

---

* The titration of insulin was carried out by E. J. Cohn, J. D. Ferry, and M. H. Blanchard; a brief summary of the results is given in Cohn & Edsall (pp. 501–2[10]). The data for the serum albumins are still unpublished; for permission to quote them, I am indebted to E. J. Cohn, L. E. Strong, W. L. Hughes, Jr., and M. H. Blanchard.

TABLE 1

IONIC GROUPS IN CERTAIN PROTEINS AS INFERRED FROM ANALYSIS AND FROM TITRATION DATA

(All figures are given in moles of group or residue per 100 kg. of protein.)

Group	Human Serum Albumin	Bovine Serum Albumin	Insulin
(a) Histidine	22.6	24.5	34.2
(b) Arginine	35.3	35.6	20.1
(c) Lysine	84.1	84.8	17.8
(d) α-Amino	12	14	47.1
(e) Total Cationic Groups from Analysis	154.0	157.9	119.2
(f) Maximum Positive Charge from Titration	144	147	100
(g) Total Aspartic Acid	78.1	79.6	51.1
(h) Total Glutamic Acid	118.3	114.9	138.1
(i) α-Carboxyl	12.8	14.3	47.1
(j) Total Carboxyl	209.2	208.8	236.3
(k) Amide NH_3	62.8	61.7	126.2
(l) Free Carboxyl $(g + h + i - k)$	146.4	147.1	110.1
(m) Tyrosine	25.7	30.4	67.9
(n) Total Free Ionic Groups from Analysis $(= e + l + m)$	327.3	336.3	297.2
(o) Free Ionic Groups from Titration between pH 1 and 12	279	281	230

of molecular weight 46,000, as against the figure of 21 derived from the combination of the lysine value with the amino nitrogen determination. These figures, as well as a consideration of the possible errors in the amino nitrogen determination on an intact protein, lead me to hold decided reservations regarding the significance of the estimated free α-amino and α-carboxyl groups. I think it is exceedingly important that independent methods of estimating these groups should be developed and should be shown to give concordant results, before any extensive inferences concerning protein structure are drawn from the data so obtained.

Unfortunately, the titrations in our laboratory have not yet been extended to include aqueous formaldehyde solutions of the protein, as Cannan and his collaborators have done for egg albumin and β-lactoglobulin. It is, thus, impossible, at present, to give reliable estimates of the total amino groups from the titration data on the serum albumins and insulin. Investigations of this point are under way.

However, it is of some interest to record the total number of free ionic groups in these proteins, as deduced from analysis, and compare it with the total range of the titration figures between pH 1 and 12. It is assumed that the free anionic groups are represented by the free carboxyls and the phenolic hydroxyl groups of tyrosine. These data, listed in Rows n and o of TABLE 1, show, of course, a considerably larger number of groups from analysis than are determined in the titration curves. Some discrepancies of this sort are to be expected, since some, at least, of the guanidine residues of arginine still carry a positive charge at pH 12. In several proteins, there is evidence that the hydroxyl groups of the tyrosine residues are still largely in the uncharged form at this pH. Different proteins differ markedly in this respect, however, to judge from the ultraviolet absorption measurements of Crammer and Neuberger.[16] In any case, the discrepancy between the values in Rows n and o of TABLE 1 is never as great as the sum of the tyrosine and arginine residues, not to mention the perhaps still questionable α-carboxyl groups. Further extension and refinement of these incomplete titration data should lead to a more searching comparison with the analytical figures and, perhaps, to further insight which may elicit new lines of investigation.

BIBLIOGRAPHY

1. **Shedlovsky, T.**
1943. Ann. N.Y. Acad. Sci. **43**: 259.
2. **Svedberg, T., & K. O. Pedersen**
1940. The Ultracentrifuge. Oxford.
3. **Oncley, J. L.**
1941. Ann. N.Y. Acad. Sci. **41**: 121.
4. **Rothen, A.**
1942. Ann. N.Y. Acad. Sci. **43**: 229.
5. **Miller, G. L., & K. J. I. Andersson**
1942. J. Biol. Chem. **144**: 459.
6. **Sjögren, B., & T. Svedberg**
1931. J. Am. Chem. Soc. **53**: 2657.
7. **Polson, A.**
1939. Kolloid Z. **87**: 149.
8. **Brohult, S.**
1940. Nova Acta Reg. Soc. Scient. Upsaliensis. Ser. IV. **12**(4): 69.
9. **Foster, J. F., & D. French**
1945. J. Am. Chem. Soc. **67**: 687.
10. **Cohn, E. J., & J. T. Edsall**
1943. Proteins, Amino Acids and Peptides. Reinhold Publishing Corp. New York.
11. **Archibald, W. J.**
1942. Ann. N.Y. Acad. Sci. **43**: 211.

12. **Fankuchen, I.**
1941. Ann. N.Y. Acad. Sci. **41**: 157.

13. **McMeekin, T. L., & R. C. Warner**
1942. J. Am. Chem. Soc. **64**: 2393.

14. **Crowfoot, D.**
1941. Chem. Rev. **28**: 215.

15. **Cannan, R. K., A. H. Palmer, & A. C. Kibrick**
1942. J. Biol. Chem. **142**: 803.

16. **Crammer, J. L., & A. Neuberger**
1943. Biochem. J. **37**: 302.

CONCLUSION

By Hans T. Clarke

College of Physicians and Surgeons, Columbia University, New York, N.Y.

For years, the principal obstacle to attempts to gain even a preliminary idea of the fine structure of proteins has been the lack of accurate knowledge of the quantitative distribution of the constituent amino acids. The first step in this direction has now been achieved in the preparation of a complete analysis of β-lactoglobulin, in which all the component amino acids are accounted for with a precision of better than 3 per cent.

As has been clearly brought out in this Conference, the determination of amino acids in proteins has involved a wide variety of approaches. It seems proper, at this time, to survey briefly the advantages and disadvantages of the various techniques here discussed.

The difficulties inherent in quantitative isolation procedures have been pointed out. These restrict the utility of the methods involving specific precipitants, which are applicable to a wide variety of individual amino acids, but are laborious and, compared to colorimetric and microbiological methods, require relatively large amounts of proteins. The solubility product method offers more encouraging possibilities and merits further exploration, with especial reference to its accuracy in the estimation of amino acids which are present as minor constituents.

The procedures involving the principles of chromatography and distribution between immiscible solvents offer considerable promise, but have not been developed to the stage of utility attained by most of the other methods under consideration.

The development of the microbiological methods discussed has been remarkably rapid. Their specificity is high, they are applicable to fantastically small amounts, and their precision is astonishingly great, particularly in view of the uncertainty which always attaches to procedures involving living systems. Although it seems unlikely that they will soon be placed upon a theoretical basis, a brilliant future can safely be predicted for these methods.

Considerable progress has recently been made in colorimetric methods, through the adaptation of photoelectric techniques, which have, moreover, made possible the development of spectrophotometry in the ultraviolet. However, the application of these methods is restricted to a relatively small number of individual amino acids. Similar conditions exist with respect to enzymatic procedures.

The isotope dilution method, which, theoretically, should yield results of particular value as absolute standards of reference for testing the validity of other analytical procedures, can at present be carried out only in the few laboratories in which a mass spectrometer is operated. If isotope analytical services, available to any research worker, are organized, the isotope dilution methods will become more generally useful. Isotopic material for the synthesis of labeled amino acids can be purchased on the open market, and the development of appropriate synthetic procedures should demand no special equipment. For the elaboration of suitable methods of isolation, the specific precipitants should render valuable aid.

The basic uncertainty which underlies all analytical work in this field, namely, the degree to which quantitative data secured with hydrolysates represent the true composition of proteins, appears to have been resolved in the case of lactoglobulin, in which not only all of the nitrogen, but all of the carbon and hydrogen, are accounted for. However, with other proteins, the occurrence, during hydrolysis, of a significant degree of racemization and even more profound decompositions must be anticipated. The extent to which such side reactions occur with amino acids *per se* can now be estimated by methods involving isotopes, but no certainty exists with regard to their occurrence with amino acids while they are still in peptide combination. It is conceivable that an approach can be made by comparative quantitative studies with acid hydrolysates of proteins, before and after treatment with proteolytic enzymes.

The next step in the general program is, clearly, the complete analysis of a wide variety of demonstrably homogeneous proteins. Methods for the preparation of such material have not yet attained the same degree of perfection as the analytical procedures. The mere fact that a protein can be obtained in crystalline condition does not, in itself, form a guarantee of homogeneity comparable to that for the simpler compounds of classical organic chemistry. In compounds of high molecular weight and complex composition, the danger that mixed crystals will be formed is always present in a high degree. Gravitational methods can indicate homogeneity of molecular dimension; electrophoretic methods indicate approximate uniformity of topological distribution of electric charges on the molecule. However, even when both of these criteria are satisfied, true chemical homogeneity remains to be assured. Certainty can, in some instances, be secured by means of the phase rule procedures developed by Northrop and Kunitz. Amino acid analysis should, theoretically, supply further information on this question, but there is obviously a need for ever-increasing accuracy and precision. Proteins which exhibit specific physiological activities, such as hormones and enzymes, lend themselves to individual methods of study. It is, perhaps, these that offer the most promise for the immediate future, in the general problem of protein constitution.

DATE DUE

DEMCO 38-297

PUBLICATIONS

OF THE

NEW YORK ACADEMY OF SCIENCES

(Lyceum of Natural History, 1817–1876)

---

(1) The **Annals** (octavo series), established in 1823, contain the scientific contributions and reports of researches, together with the records of meetings of the Academy. The articles which comprise each volume are printed separately, each in its own cover, and are distributed immediately upon publication. The price of the separate articles depends upon their length and the number of illustrations, and may be ascertained upon application to the Executive Secretary of the Academy.

(2) The **Transactions,** reestablished as Series II in 1938, contain extended abstracts of papers presented before the regular sectional meetings of the Academy and other matters of general interest to Members, and are published monthly from November to June, inclusive.

Current numbers of the Annals and Transactions are sent free to all Members of the Academy desiring them.

(3) The **Special Publications,** established in 1939, are issued at irregular intervals as cloth-bound volumes. The price of each volume will be advertised at time of issue.

(4) The **Memoirs** (quarto series), established in 1895, are issued at irregular intervals. It is intended that each volume shall be devoted to monographs relating to some particular department of Science. Volume I, Part 1, is devoted to Astronomical Memoirs, Volume II to Zoological Memoirs. No more parts of the Memoirs have been published to date. The price is one dollar per part.

(5) **The Scientific Survey of Porto Rico and the Virgin Islands** (octavo series), established in 1919, gives the detailed reports of the anthropological, botanical, geological, paleontological, zoological, and meteorological surveys of these islands.

Subscriptions and inquiries concerning current and back numbers of any of the publications of the Academy should be addressed to

EXECUTIVE SECRETARY

The New York Academy of Sciences
Central Park West at Seventy-ninth Street
New York, N.Y.